



PCT/CN03/00580

# 证 明

REC'D 04 SEP 2003

WIPO PCT

本证明之附件是向本局提交的下列专利申请副本

申 请 日： 2003 06 30

申 请 号： 03 1 48529.4

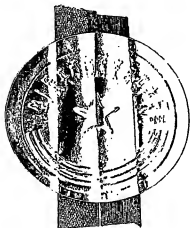
申 请 类 别： 发明

发明创造名称： 一种基于 D N A 芯片的基因分型方法及其应用

申 请 人： 清华大学；北京博奥生物芯片有限责任公司

发明人或设计人： 高华方；马雪梅；张翅；陈谦；周一鸣；王栋；章轶哲；  
田玉娥；张蕊；兰更欣；周玉祥；程京

BEST AVAILABLE COPY



**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

中华人民共和国  
国家知识产权局局长

王景川

2003 年 8 月 11 日



## 权利要求书

4

1、一种对目的基因进行分型的方法，包括：

a) 从合适样品中分离含有目的基因的靶细胞；

b) 目的核苷酸序列的制备，所述目的核苷酸序列至少是所述目的基因的一部分；  
与所述目的基因不相关的另一核苷酸序列的制备；

c) 提供一种芯片，所述芯片上固定了能和所述目的核苷酸序列杂交的寡核苷酸探针和至少一种以下对照探针：阳性对照探针、阴性对照探针、杂交对照探针和固定化对照探针；

d) 使步骤 b) 中所述目的核苷酸序列和/或所述另一核苷酸序列和步骤 c) 中所提供的芯片杂交，根据所述目的核苷酸序列和/或所述另一核苷酸序列与所述芯片上的所述探针的杂交结果来对所述目的基因进行分型。

2、根据权利要求 1 所述的方法，其特征在于：所述靶细胞是白细胞。

3、根据权利要求 1 所述的方法，其特征在于：所述目的基因是人类白细胞抗原基因。

4、根据权利要求 1 所述的方法，其特征在于：所述合适样品选自血液、唾液、毛发、含有核酸的人类组织和其它任何含有有核细胞的人类组织。

5、根据权利要求 4 所述的方法，其特征在于：所述血样选自血清、血浆和全血。

6、根据权利要求 5 所述的方法，其特征在于：所述血样是新鲜的或低温保存的。

7、根据权利要求 1 所述的方法，其特征在于：所述靶细胞是用微型磁珠从合适样品中分离出来的。

8、根据权利要求 7 所述的方法，其特征在于：所述微型磁珠的直径为 5—200  $\mu\text{m}$ 。

9、根据权利要求 1 所述的方法，其特征在于：所述目的核酸序列的制备包括核酸扩增步骤。

10、根据权利要求 9 所述的方法，其特征在于：所述目的核酸序列是通过对所分离的靶细胞直接进行核酸扩增得到的。

11、根据权利要求 9 所述的方法，其特征在于：所述目的核苷酸序列是通过以所述分离的靶细胞的核酸为模板进行核酸扩增得到的。

12、根据权利要求 9 所述的方法，其特征在于：所述核酸扩增步骤为聚合酶链式反应、连接酶链式反应、基于核酸序列的扩增、链置换扩增、滚环复制或转录介导的扩增。



13、根据权利要求 12 所述的方法，其特征在于：所述转录介导的扩增是由 T7 启动子启动的。

14、根据权利要求 12 所述的方法，其特征在于：所述聚合酶链式反应为不对称 PCR。

15、根据权利要求 14 所述的方法，其特征在于：所述不对称 PCR 的上下游引物比例为 1:5 — 1:200。

16、根据权利要求 14 所述的方法，其特征在于：所述不对称 PCR 的上下游引物的 Tm 值相同或不同。

17、根据权利要求 14 所述的方法，其特征在于：所述不对称 PCR 的上下游引物 Tm 值相差 1—20℃。

18、根据权利要求 14 所述的方法，其特征在于：所述不对称 PCR 有三条引物，其中两条引物的 Tm 值相近，第三条的 Tm 值与所述两条引物的 Tm 值相差 1—20℃。

19、根据权利要求 14 所述的方法，其特征在于：所述不对称 PCR 的引物是直链引物或包含发卡结构的引物。

20、根据权利要求 12 所述的方法，其特征在于：所述聚合酶链式反应的退火温度为单个或多个。

21、根据权利要求 20 所述的方法，其特征在于：所述退火温度之间相差 1—20℃。

22、根据权利要求 1 所述的方法，其特征在于：所述步骤 b) 中目的核酸序列为单链 DNA 或 RNA。

23、根据权利要求 22 所述的方法，其特征在于：所述单链 DNA 或 RNA 为正链，或负链。

24、根据权利要求 1 所述的方法，其特征在于：所述步骤 b) 中得到的目的核酸序列是被标记的。

25、根据权利要求 24 所述的方法，其特征在于：所述标记为荧光素或生物素标记。

26、根据权利要求 1 所述的方法，其特征在于：所述另一核苷酸序列是与固定在所述芯片上的阳性对照探针、阴性对照探针或杂交对照探针互补的基因序列。

27、根据权利要求 1 所述的方法，其特征在于：所述固定于芯片上的探针为正链或负链探针。

28、根据权利要求 1 所述的方法，其特征在于：所述固定于芯片上的探针是被修饰过的。



29、根据权利要求 28 所述的方法，其特征在于：所述探针修饰选自 5' -NH<sub>2</sub> 修饰，5' -SH 修饰，5' -polyT、5' -poly A、5' -poly C 或 5' -poly G 修饰，5' -生物素修饰，3' -NH<sub>2</sub> 修饰，3' -SH 修饰，3' -polyT、3' -poly A、3' -poly C 或 3' -poly G 修饰和 3' -生物素修饰。

30、根据权利要求 1 所述的方法，其特征在于：所述芯片含有 1-400 种不同类型的探针。

31、根据权利要求 1 所述的方法，其特征在于：所述芯片上有复合探针阵列，每一个阵列包括 1-400 种不同类型的探针。

32、根据权利要求 1 所述的方法，其特征在于：所述探针是被固定到芯片上的，固定化温度为 37-100 °C。

33、根据权利要求 1 所述的方法，其特征在于：所述芯片是经过修饰的。

34、根据权利要求 33 所述的方法，其特征在于：所述芯片修饰为在芯片表面包被以下任意一种官能团：醛基，氨基，聚赖氨酸，巯基，牛血清蛋白，链霉亲和素，琼脂糖凝胶和聚丙烯酰胺胶。

35、根据权利要求 1 所述的方法，其特征在于：所述探针的序列正确性、纯度、末端修饰情况是经过检验的。

36、根据权利要求 35 所述的方法，其特征在于：所述探针的序列正确性、纯度、末端修饰情况的检验通过 DHPLC 完成。

37、根据权利要求 1 所述的方法，其特征在于：所述芯片上固定有一种探针的多个拷贝。

38、根据权利要求 1 所述的方法，其特征在于：所述芯片上固定的一种探针的拷贝数为 1-10。

39、根据权利要求 1 所述的方法，其特征在于：所述一种探针的多个拷贝连续或分开固定于所述芯片上。

40、根据权利要求 39 所述的方法，其特征在于：一种阳性对照探针的多个拷贝固定于所述芯片上；当所述固定的阳性对照探针与步骤 b) 所述目的核苷酸序列或另一核苷酸序列杂交时，所述固定的阳性对照探针序列和长度的变化，可产生从强到弱或从弱到强变化的信号。

41、根据权利要求 1 所述的方法，其特征在于：所述阳性对照探针与目的核苷酸序列、以目的核苷酸序列为模板扩增产生的核苷酸序列或人工合成的核苷酸序列的一部分互补。



42、根据权利要求 41 所述的方法，其特征在于：阴性对照探针与所述阳性对照探针有 1—3 个碱基对的错配。

43、根据权利要求 1 所述的方法，其特征在于：所述杂交对照探针与目的基因无同源性的人工合成的核苷酸序列互补。

44、根据权利要求 43 所述的方法，其特征在于：所述杂交对照探针与人工合成的标记的核苷酸序列完全互补，或者与人工合成的标记的核苷酸序列有 1—2 个碱基对的错配。

45、根据权利要求 1 所述的方法，其特征在于：所述固定化对照探针不会与目的核苷酸序列发生反应而产生杂交信号。

46、根据权利要求 1 所述的方法，其特征在于：所述固定化对照探针的一端进行了化学修饰，另一端有一种可被检测的标记。

47、根据权利要求 1 所述的方法，其特征在于：所述芯片上包含了一种阳性对照探针、一种阴性对照探针、一种杂交对照探针和一种固定化对照探针。

48、根据权利要求 1 所述的方法，其特征在于：所述阳性对照探针、阴性对照探针、杂交对照探针和固定化对照探针固定在所述芯片的四个角上、中间或按其它合适的方式排列或随机分布在探针阵列中。

49、根据权利要求 1 所述的方法，其特征在于：所述步骤 d) 中的杂交反应在含有氯化钠和柠檬酸钠和表面活性剂的缓冲液中进行。

50、根据权利要求 49 所述的方法，其特征在于：所述缓冲液为  $3\times\sim 10\times$  的 SSC。

51、根据权利要求 49 所述的方法，其特征在于：所述表面活性剂的质量浓度从 0.05% 到 5%。

52、根据权利要求 49 所述的方法，其特征在于：所述表面活性剂选自十二烷基磺酸钠、Triton 100 和 十二烷基肌氨酸钠。

53、根据权利要求 1 所述的方法，其特征在于：所述步骤 d) 中的杂交反应温度范围为  $42^{\circ}\text{C}$  到  $70^{\circ}\text{C}$ 。

54、根据权利要求 1 所述的方法，其特征在于：所述方法包括在杂交反应结束后进行清洗的步骤。

55、根据权利要求 54 所述的方法，其特征在于：所述清洗用缓冲液含有质量浓度为 0% 到 2% 的表面活性剂。

56、根据权利要求 54 所述的方法，其特征在于：所述清洗进行 5 分钟到 30 分钟。

57、根据权利要求 1 所述的方法，其特征在于：所述方法中固定化效率是通过固定化对照探针的信号值来评价的。



58、根据权利要求 1 所述的方法，其特征在于：所述方法中总体杂交效率是通过分析杂交对照探针和与目的基因无关的合成的带标记的核苷酸序列的杂交信号来评价的。

59、根据权利要求 1 所述的方法，其特征在于：所述方法中杂交特异性通过分析阳性对照探针的杂交信号和阴性对照探针的杂交信号的比值以及阳性杂交对照探针的杂交信号和阴性杂交对照探针的杂交信号的比值来进行评价，比值越高则说明特异性越好。

60、根据权利要求 1 所述的方法，其特征在于：在涉及一组紧密相关的探针杂交中，阳性杂交信号的判断标准是：

- a) 杂交信号和背景噪音的比值大于 3；
- b) 杂交信号与相应阳性对照探针的杂交信号的比值在一个预先确定的范围内；
- c) 比较所有根据步骤 b 和 c 产生阳性杂交信号的所有探针的杂交信号，或者仅一个探针表现出阳性杂交信号时比较表现出最强杂交信号的两个探针的杂交信号，来确定杂交信号是阴性还是阳性。
- d) 在紧密相关的探针组里只有 2 个或少于 2 个阳性信号。

61、根据权利要求 1 所述的方法，其特征在于：所述寡核苷酸探针和 HLA 目的基因互补。

62、根据权利要求 61 所述的方法，其特征在于：所述寡核苷酸探针含有如下的序列：

- a) 能够在高严谨性条件下与目的 HLA 核苷酸序列或其互补链进行杂交的序列；所述序列选自序列表中的序列 1—214；
- b) 能够与所述目的 HLA 核苷酸序列至少有 90%匹配率的序列，所述序列为正义链或负义链；所述序列为序列表中的序列。

63、根据权利要求 61 所述的方法，其特征在于：所述寡核苷酸探针是序列表中的序列或其互补序列。

64、根据权利要求 61 所述的方法，其特征在于：所述芯片上包括序列表中的所有寡核苷酸序列或其互补序列。

65、用于对 HLA 基因进行分型的寡核苷酸探针包含如下序列：

- a) 能够在高严谨杂交条件下与目的 HLA 基因核苷酸序列或其互补链进行杂交的序列；所述序列选自序列表中的序列 1—214；
- b) 能够与目的 HLA 核苷酸序列或其互补序列至少有 90%匹配率的序列，所述序列选自序列表中的序列 1—214。



66、根据权利要求 65 所述的方法，其特征在于：所述寡核苷酸探针是 DNA、RNA、PNA 或其它衍生物。

67、根据权利要求 65 所述的方法，其特征在于：所述寡核苷酸探针是序列表中的序列，或其互补序列。

68、根据权利要求 65 所述的方法，其特征在于：所述寡核苷酸探针是被标记的。

69、根据权利要求 68 所述的方法，其特征在于：所述标记选自化学标记，酶标记，免疫标记，放射性标记，荧光标记、化学发光标记和荧光共振能量转移标记。

70、一种对 HLA 目的基因进行分型的寡核苷酸探针阵列，包括可以固定大量寡核苷酸探针并适于杂交反应的固相载体，所述探针中至少一条探针具有如下序列：

a) 能够在高严谨性条件下与目的 HLA 核苷酸序列或其互补链进行杂交的序列，所述序列为序列表中的序列；

b) 能够与目的 HLA 核苷酸序列或其互补序列至少有 90%匹配的序列，所述序列为序列表中的序列。

71、根据权利要求 70 所述的阵列，其特征在于：所述寡核苷酸探针包括 DNA、RNA、PNA 或其衍生物。

72、根据权利要求 70 所述的阵列，其特征在于：所述寡核苷酸探针中至少一条探针含有表 1 中的序列或其互补序列。

73、根据权利要求 70 所述的阵列，其特征在于：所述阵列包含序列表中所有序列或其互补序列的探针。

74、根据权利要求 73 所述的阵列，其特征在于：所述探针中至少一条是被标记的。

75、根据权利要求 74 所述的阵列，其特征在于：所述标记选自化学标记，酶标记，免疫标记，放射性标记，荧光标记、化学发光标记和荧光共振能量转移标记。

76、根据权利要求 70 所述的阵列，其特征在于：所述阵列的载体表面为硅片、塑料、玻璃、陶瓷、橡胶合或聚合物。



05-08-00

10

## 说明书

### 一种基于 DNA 芯片的基因分型方法及其应用

#### 技术领域

本发明涉及基因分析领域中一种基因分型方法及其应用，特别涉及一种基于 DNA 芯片的基因分型方法及其应用。

#### 背景技术

在不同个体之间进行组织器官移植时，受体和供体双方的相互接受程度被称为组织相容性 (histocompatibility)。导致对移植器官产生排斥反应的抗原称为移植抗原 (transplantation antigen) 或称组织相容性抗原 (histocompatibility antigen)。机体内与排斥反应有关的抗原系统多达 20 种以上，其中能引起强急性排斥反应的称为主要组织相容性抗原 (major histocompatibility antigen)，其编码基因是一组紧密连锁的基因群，称为主要组织相容性复合体 (major histocompatibility complex, MHC)。现已证明，控制机体免疫应答能力与调节功能的基因 (immune response gene, IR gene) 也存在于 MHC 内。因此，MHC 不仅与移植排斥反应有关，也广泛参与免疫应答的诱导与调节。不同种属的哺乳类动物其 MHC 及编码的抗原系统有不同的命名，小鼠的主要组织相容性抗原系统称为 H-2 系统，人类的主要组织相容性抗原系统命名为 HLA，即人白细胞抗原 (human leukocyte antigen)，是由 HLA 基因复合体 (HLA gene complex) 编码的。HLA 基因复合体是人类主要组织相容性复合体 (major histocompatibility complex, MHC) 的一部分，是一个与组织抗原和免疫应答相关的基因簇。HLA 基因复合体位于人第 6 号染色体短臂上 4000kb 的 DNA 片段内，约占人体整个基因组的 1/3000，是目前已知的人类最复杂的基因系统之一。HLA 复合体已经发现基因座位有 224 个，传统上按其产物的结构、表达方式、组织分布与功能可将这些基因座分为三类：即 HLA-I，HLA-II，HLA-III。而每一类基因座位中又有数百种等位基因 (allele)。

目前，同种异体间的组织器官移植 (allograft) 已经成为医学临床上治疗某些恶性疾病的有效手段。而异体移植 (如肾脏、心脏、皮肤等) 能否在接收者体内存活并行使正常生理功能，是临床医学的重大研究课题，也是人们极为关心的问题。免疫学研究的发展阐明了移植物的排斥反应是机体免疫系统的正常功能——排斥异物所造成的。通过众多免疫学家的深入研究，终于发现由组织相容性复合体 (MHC) 基因编码的蛋白质——人白细胞抗原 (HLA) 是移植排斥产生的主要决定因素。进一步的研究表明，HLA 在抗原递呈和免疫应答调控方面还具有极为重要的功能，对研究自





05:06:30

11

身免疫性疾病、病毒感染性疾病以及肿瘤的发病机理可提供有益的帮助 (So-loski MJ, Szerperka ME, Davies A, Wooden SL. Host immune response to intracellular bacteria: A role for MHC-linked class-IIb antigen-presenting molecules. Proc.Soc.Exp.Biol.Med. 2000;224:231-9.)。由于HLA在群体中不同个体之间呈现高度多态性,不同个体对抗原的免疫应答程度和能力也存在差异,因而对疾病的易感性和抗性也有所不同,因此研究HLA对诊断和防治疾病具有积极的意义。另外,由于HLA具有单倍型遗传和共显性表达的特点,所以HLA还可作为法医鉴定、亲子鉴定及遗传学研究的重要指标。但是,这些工作的开展都必须依赖于对HLA进行准确地分型。因此,HLA分型技术是现代医学和免疫学研究的支撑性技术。

HLA分型指的是用血清学、细胞学、分子生物学的技术和方法或其它的技术和方法对人类的HLA抗原特异性进行准确鉴定的一种技术。目前常用的鉴定方法有三种,它们分别是:60年代建立起来的并不断完善的血清学分型法、细胞学分型法和80年代后期建立起来的并快速发展的DNA分型法。

### 1、血清学分型技术

血清学分型法有白细胞凝集试验、淋巴细胞毒试验、抗体依赖细胞介导淋巴细胞毒试验等。目前普遍采用的是Terasaki于1964年建立,后经几度改良的微量淋巴细胞毒试验(microlymphocytotoxicity test)或称补体依赖的细胞毒试验(complement dependent cytotoxicity test)。基本原理是抗体(IgG或IgM)与活的淋巴细胞膜上相应的抗原结合,在补体存在的条件下破坏淋巴细胞膜,染料(曙红或锥蓝)进入细胞而染色。活细胞因细胞膜通透性没有改变,染料不能进入细胞而不能着色。通过显微镜观察,估算出死亡细胞的数目,用计分方法反映细胞毒反应强度。一般必须通过两个分型血清交叉进行。根据分型血清的特异性就可以检出被检细胞的抗原特异性。标准抗血清多取自多次经产妇产或计划免疫志愿者。标准血清学方法存在着诸多的不足,例如:①随着新的等位基因特异性被发现,血清学方法难以获得相应的标准抗血清;②较多的交叉反应使亚型分辨困难;③标准血清的筛选技术复杂、难度大,至今C位点的血清学分型仍缺乏理想的试剂。

80年代末期,出现了单维等电聚焦(ID-IEF)法用于HLA抗原的分型,但其技术条件高、操作复杂,只能作为血清学的一种补充。基于血清学原理和分辨水平的单克隆抗体分型技术于1994年首先由美国莱姆德公司研制成功,克服了许多血清学的不足,分型时间缩短、特异性提高,已被很多临床实验室采纳。虽然目前血清学分型仍是一种占有重要地位的分型技术,但是由于其具有一些在技术上无法克服的缺陷,近年来国际上已经把研究重点放到基因分型上。



05-05-00

11

## 2、细胞学分型技术

细胞学分型是采用混合淋巴细胞培养(mixed lymphocyte culture, MLC)的方法。当同种异体淋巴细胞的 HLA 抗原不同时,将两种淋巴细胞混合培养,将会刺激对方细胞发生增殖反应,通过检测淋巴细胞的增殖程度就可了解两者抗原的同源性程度。HLA-Dw 特异性与 HLA-DP 特异性可分别通过纯合分型细胞(homozygote typing cell, HTC)及预致敏淋巴细胞试验(primed lymphocyte test, PLT)检测。二种方法的基本原理均是判断淋巴细胞在识别非己 HLA 抗原决定簇后发生的增殖反应。由于分型细胞来源困难以及操作手续繁琐,细胞学分型技术已逐渐被淘汰。

## 3、HLA 的 DNA 分型技术

HLA 的 DNA 分型技术始于 80 年代末,近年逐渐发展成熟。主要方法包括限制性酶切片段长度多态性分析(RFLP)、测序与单链构象多态性分析(PCR-SSCP)、寡核苷酸探针杂交(PCR-SSOP)和顺序特异引物聚合酶链反应(PCR-SSP)方法。尽管血清学对 I 类抗原分型有较高的准确性和实用性,但与 DNA 分型结果比较仍存在一定的误差。在 HLA-I 类血清学误差中,1/3 是亚型误差、50%是由于血清学空白位点经 DNA 证实存在第二个位点。由于 HLA 的基因分型法准确率高,操作相对简单,近年来已被国内外大多数实验室采用。

### (1) 限制性酶切片段长度多态性(RFLP)检测技术

这是于 80 年代后期首先建立的对多态性进行检测的 DNA 分析技术(Oertel M, Berr F, Schroder S, Schwarz R, Tannapfel A, Wenzke M et al. Acute rejection of hepatic allografts from HLA-DR13 (Allele DRB1(\*)1301)-positive donors. Liver Transpl. 2000;6:728-33.) (Erlich U.S. Pat. No. 4,582,788)。个体间抗原特异性来自氨基酸顺序的差别,后者由编码基因的碱基顺序不同所决定。这种碱基顺序的差别造成限制性内切酶识别位置及酶切位点数目的不同,从而产生数量和长度不一的 DNA 酶切片段。用特异性探针针对整个基因组 DNA 酶切片段进行杂交,即可分析限制性长度片段多态性(restriction fragment length polymorphism, RFLP)。一定的内切酶组合所得到的 HLA-RFLP 可以和传统方法测定的 HLA 特异性型别相关。80 年代末发展起来的 PCR (polymerase chain reaction) 技术已被用于 RFLP 分析,即用等位特异限制酶裂解 PCR 扩增的片段,然后再进行分析,从而大大提高了灵敏度。

其缺点是获得一个完全的 HLA-II 样品分型至少需要 2 周的时间,且探针的放射性标记使得该技术只能在少数实验室中进行。

### (2) PCR/SSO 技术

到 1989~1991 年,分子生物学已经建立起新的 HLA-II 类基因分型标准(Middleton



07-08-10

3

D, Williams F, Cullen C, Mallon E. Modification of an HLA-B PCR-SSOP typing system leading to improved allele determination. *Tissue Antigens* 1995;45:232-6.)。此法用人工合成的 HLA 型别特异的寡核苷酸序列作为探针,与待检细胞经 PCR 扩增的 HLA 基因片段杂交,从而确定 HLA 型别,PCR 技术可将 HLA 复合体上指定基因片段特异性地扩增 5~6 个数量级;而专门设计的序列特异的寡核苷酸(sequence specific oligonucleotide, SSO)探针又能探测出等位基因间 1~2 个核苷酸的差异,故 PCR/SSO 技术具有灵敏度高、特异性强、需样本量少等优点。

该方法的缺点是:对每一种探针都要制备条件一致的膜,随着探针数目的增加,膜的数目可能会很大。如果想使膜重复使用,则必须在杂交完成后将探针清除掉,而这种工作目前还不能实现自动化,可以想象该工作的复杂程度。

鉴于该方法的不足,现在世界上一些有实力的公司正在对其进行改进,如美国的 Roche 公司已经于最近开发出了一种反向杂交技术 (Reverse line strip typing method),其显色方法是酶学方法,其操作程序非常复杂,要求极为严格,操作不当很容易得不到理想的结果。

### 3.3、PCR/SSP 技术

目前常规的 HLA-DNA 分型技术,包括上述的 PCR/RFLP、PCR/SSO 等,最终均需标记的特性探针与扩增产物进行杂交,再分析结果。而 PCR/SSP 方法是设计出一整套等位基因组特异性引物 (sequence specific primer, SSP),借助 PCR 技术获得 HLA 型别特异的扩增产物,可通过电泳直接分析带型决定 HLA 型别,从而大大简化了实验步骤,使整个检测过程缩短为 2~3 个小时 (Mytilineos J, Lempert M, Scherer S, Schwarz V, Opelz G. Comparison of serological and DNA PCR-SSP typing results for HLA-A and HLA-B in 421 Black individuals: a Collaborative Transplant Study report. *Hum.Immunol.* 1998;59:512-7.)。

但该方法的缺点也是显而易见的,对于一份未知样品,需要用各种序列特异的引物去试,需要做大量的 PCR 实验,还要做大量的琼脂糖凝胶电泳,工作量是巨大的。但是该方法快速简便,所以已经有数家国际大公司开发出了试剂盒和自动化分析仪器,使工作量大大降低,使其成为目前国际分型市场上的一种重要方法。但是,由于 HLA 序列的高度多态性和特异性,对同一种型号下的各种亚型的区分就比较困难,很难做到较高的分辨率。

由于传统方法在 II 类抗原分型方面困难较大,故大部分的基因分析型方法目前主要用于 II 类基因座。此外,目前已建立的 HLA 基因分型技术还包括 PCR 单链构像多态性分析 (PCR-single strand conformational polymorphism, PCR-SSCP) 和 PCR 异



源二聚体电泳多态即 PCR 指纹图 (PCR fingerprinting) 分析。DNA 分型技术的应用, 使 HLA 型别分析达到了更精细的水平, 并因此发现了更多的 HLA 多态性。HLA 的 DNA 分型技术现已成为血清学方法的有力竞争者, 并可能在不久的将来完全取而代之。

#### 4、基于 DNA 芯片的 HLA 分型技术

鉴于 HLA 分型的在医学、免疫学、遗传学等方面的重要性, HLA 分型技术一直是一个倍受关注的研究热点。除了已有的分型技术如血清学方法、细胞学方法、DNA 分型方法得到进一步发展外, 近年来又有一些新的方法和技术出现。例如, Roche 公司的 Sturniolo 等根据人类白细胞抗原空间构像的特异性, 利用噬菌体展示技术和蛋白质芯片技术相结合而发展了一种可用于 HLA 分型的新技术; Worral 等报道引物延伸技术和时间飞行质谱相结合有可能发展成为一种高通量的 HLA 分型技术, 只是目前其成本还较高; 近年来快速发展的毛细管电泳技术有可能成为另一种对 HLA 进行高分辨率分型的方法。另外, 随着生物芯片技术的迅速发展, 基于生物芯片的 HLA 分型技术正日益受到人们的关注。

自从 Fodor 等 1991 年在著名杂志 Science 上提出 DNA 芯片的概念后, 近年来以 DNA 芯片为代表的生物芯片 (biochip) 技术得到了迅猛发展, 目前已有多种不同功用的芯片问世, 而且, 这些芯片中有的已经在生命科学研究中开始发挥重要作用。生物芯片采用了光刻加工等缩微技术, 将生命科学中许多不连续的过程如样品制备、化学反应和检测等步骤移植到芯片中并使其连续化和微型化。这种基于微加工技术发展起来的生物芯片, 可以把成千上万乃至几十万个生命信息集成在一个很小的芯片上, 对基因、抗原和活体细胞等进行测试分析。用这些生物芯片所制作的各种不同用途的生化分析仪和传统仪器相比较具有体积小、重量轻、成本低, 便于携带, 防污染、分析过程自动化、分析速度快、所需样品和试剂少等诸多优点。因此, 生物芯片技术必将在 HLA 分型领域发挥革命性的作用。虽然目前已经开发出了各种 HLA 分型用的试剂盒和方法 (Kashiwase K. DNA typing methods of HLA. Rinsho Byori 1999; Suppl 110:99-106.), 但由于 HLA 系统是迄今为止发现的最为复杂的人类基因序列, 而且每年都有 5% 左右的新序列被发现, 所以 HLA 分型的难度将不断增加。而生物芯片所具有的高通量、并行分析、易于操纵、所需试剂量少、检测快速等特点使其成为 HLA 分型最为理想的工具, 现在已经有一些研究者开展了这方面的积极探索 (Cao K, Chopek M, Fernandez-Vina MA. High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). Rev. Immunogenet. 1999; 1:177-208.; Guo Z, Hood L, Petersdorf EW. Oligonucleotide arrays for high resolution HLA typing. Rev. Immunogenet. 1999; 1:220-30.)



## 发明创造内容

本发明的目的是提供一种对目的基因进行分型的方法。

本发明所提供的对目的基因进行分型的方法包括：

a) 从合适样品中分离含有目的基因的靶细胞；

b) 目的核苷酸序列的制备, 所述目的核苷酸序列至少是所述目的基因的一部分；  
与所述目的基因不相关的另一核苷酸序列的制备；

c) 提供一种芯片, 所述芯片上固定了能和所述目的核苷酸序列杂交的寡核苷酸探针和至少一种以下对照探针：阳性对照探针、阴性对照探针、杂交对照探针和固定化对照探针；

d) 使步骤 b) 中所述目的核苷酸序列和/或所述另一核苷酸序列和步骤 c) 中所提供的芯片杂交, 根据所述目的核苷酸序列和/或所述另一核苷酸序列与所述芯片上的所述探针的杂交结果来对所述目的基因进行分型。

此方法适用于对靶细胞（如白细胞）的目的基因的进行分型。其它的靶细胞还包括动物细胞、植物细胞、真菌细胞、细菌细胞、重组细胞、培养细胞等。

此方法可用于任何目的基因的分型, 如人类白细胞抗原 (HLA) 基因, HLA-I 类基因或 II 类基因。

任何合适的样品, 如含有人类核酸的血液、唾液、毛发、人体组织等都能用于此方法。血液样品指血清、血浆或全血, 可以是新鲜的或者冷藏的。

靶细胞能用任何合适的方法从样品中被分离出来。例如, 靶细胞能用磁珠从样品中分离。这种磁珠直径最好为 5 $\mu$ m 到大约 200 $\mu$ m。

这种磁珠能用任何适当的方法制备。例如, 专利 CN 01/109870.8、专利 CN Pat. NO. 01134861.5 或 WO02/075309 中公布的方法。

任何适当的磁性物质都可用此方法制备磁性微球。这些磁性物质包括铁磁物质, 亚铁磁物质, 顺磁物质以及超顺磁物质等。构成磁珠的顺磁物质可以是顺磁金属氧化物。顺磁金属氧化物优选的是某种过渡元素或它们的合金的金属氧化物, 如：铁、铜、钴、锰、镍、钽、锌、锆等任何合适的过渡元素或  $\text{Fe}_3\text{O}_4$  或  $\text{Fe}_2\text{O}_3$ 。磁珠中的磁性物质可包含金属成份。金属成份可以是过渡元素或它们的合金, 如铁、铜、钴、锰、镍、钽、锌、锆和钴-钽-锆合金等。

磁珠可由可用的前体小珠制成, 也可用原材料制备, 或者由专利 U. S. Patent No. 5,834,121 中公布的单体交联形成的包被于金属氧化物外形成的刚性聚集体制成。可作为可磁化材料的合适物质包括：磁铁矿、铁氧体、锰、钴、镍、赤铁矿以及各种合金等铁的氧化物。磁铁是最适合的金属氧化物。通常, 金属盐被转化为金属氧化物,



然后用聚合物包被,或被含有还原基团的热塑树脂小珠吸收。如果用金属氧化物颗粒制备疏水的前体颗粒,必须用热塑乙烯聚合体包被。最好用聚苯乙烯,它能与多孔基质结合。磁性颗粒能用以下几种方法制备,参见磁学和磁性材料杂志, 15-18:1117-18, 1980; Matijevic, Acc. Chem. Res., 14:22-29, 1981 以及 U.S. Patent. No. 5091206; 4774265; 4554088; and 4421660。前体小珠的制备方法见专利 U.S. Patent. Nos. 5395688; 5318797; 5283079; 52327892; 5091206; 4965007; 4774265; 4654267; 4490436; 4336173; and 4421660, 或者使用合适的商用产品。前体小珠最好是由聚乙烯包被的磁性小珠。这样的磁珠可以从 Dynal 公司 (Lake Success, N. Y.), Rhone Poulenc (France) 公司和 SINIEF (Trondheim, Norway) 获得。也可以使用有色小珠或磁珠,其表面先包被一层不稳定的聚合物,其外层再包被一层稳定的刚性聚合物。

所述目的核苷酸序列的制备可以包括扩增的步骤。可用分离的靶细胞直接扩增,也可以先提取 DNA 再扩增。如可利用磁珠从全血中分离白细胞,用提取的核酸或直接用分离的白细胞作模板,扩增目的核酸序列,扩增得到的单链 DNA 或 RNA 可含有荧光或生物素标记的,标记的 DNA 或 RNA 可不经纯化直接用于杂交。

可使用任何适当的扩增方法,如:聚合酶链反应 (polymerase chain reaction, PCR), 连接酶链反应 (ligase chain reaction, LCR), 基于核酸序列的扩增 (nucleic acid sequence-based amplification, NASBA), 链置换扩增 (strand displacement amplification, SDA)、滚环复制 (RCA) 和转录介导的扩增 (transcription mediated amplification, TMA) 等。TMA 方法最好使用 T7 启动子。

单链 RNA 也可以通过 TMA 方法得到。含有一个 T7 启动子的引物使用 RNA 聚合酶用于扩增。用单链 DNA 或 RNA 杂交,避免了用双链所带来的 PCR 产物的纯化和变性性以及杂交信号弱或丢失的问题。

最好使用不对称 PCR 进行样品制备。不对称 PCR 中两条引物可以是任何适当的比例,比如从 1:5 到 1:200。不对称 PCR 中两条引物可以有相同的或者不同的  $T_m$  值,例如两引物的  $T_m$  值的差异可以是  $1^\circ\text{C}$  到  $20^\circ\text{C}$ 。不对称 PCR 也可以使用三条引物,两条有相同或相近的  $T_m$ , 另一条相差范围  $1^\circ\text{C}$  到  $20^\circ\text{C}$ 。引物可以是直链或带有发卡结构。可使用单一或组合的退火温度。例如:不同的退火温度可以相差  $1^\circ\text{C}$  到  $20^\circ\text{C}$ 。有较低的  $T_m$  值的引物可以用于双链的扩增,另一条有较高的  $T_m$  值的引物可以在有一定双链后扩增单链 DNA。通过这种 PCR 过程产生的单链目的核苷酸可以不经纯化直接用于杂交。

所述制备目的核酸序列的方法可以用于核酸样品的快速制备和对样品进行主动式操作以便构建微全分析系统 (或芯片实验室)。



步骤 b) 制备的目的核苷酸序列可以是单链、双链或三链。步骤 b) 制备的目的核苷酸序列最好是单链 DNA 或 RNA。步骤 b) 制备的目的核苷酸序列可以是正义链或反义链。单链 DNA 或 RNA 最好是正义链或反义链。步骤 b) 也可以制备带标记的目的核苷酸序列, 最好使用荧光或生物素标记。

所述另一寡核苷酸序列最好与芯片上的阳性对照探针、阴性对照探针或杂交对照探针互补。

所述固定于芯片上的探针可以是正义链或反义链。这些芯片上的探针可以被修饰, 修饰方法可以是 5' -NH<sub>2</sub> 修饰、5' -SH 修饰、5' -polyT(or A, C or G) 修饰、5' -biotin 修饰、3' -NH<sub>2</sub> 修饰、3' -SH 修饰、3' -polyT(or A, C or G) 修饰和 3' -biotin 修饰等。

芯片上可固定任何合适数目和类型的探针。例如: 芯片能有 1—400 种不同类型的探针。再比如: 芯片能含有多个探针阵列, 每一个阵列可以含有 1—400 种不同类型的探针。

所有探针可在任何适当的温度下固定于芯片上。如 37℃~100℃。芯片表面可以被修饰, 典型的修饰包括醛基、氨基、聚赖氨酸、巯基、BSA、链亲和素、琼脂糖凝胶和聚丙烯酰胺胶修饰。

可以对所述探针的序列、纯度、末端标记情况等进行评估。最好是使用 DHPLC。如通过 DHPLC 可测得探针的纯度和末端修饰状况。末端修饰是探针质量的一个重要的指标。DHPLC 能够分离末端修饰的探针和没被修饰的探针, 这样可以用于探针的质量控制。一个用 DHPLC 来估测探针末端 NH<sub>2</sub> 修饰的例子如图 6A—图 6D 所示, 图 6A 代表的是纯度很高的探针; 图 6B 代表的是含少量杂质的探针; 图 6C 代表的是含较多杂质的探针; 图 6D 代表的是纯度很差的探针。DHPLC 也可以用来检测探针中碱基的突变或者缺失。

芯片上每种探针可固定任何数目的拷贝, 如 1—10 拷贝。

所述多拷贝或任何合适数目探针可以以任何适当的组合形式固定于芯片。如所述多拷贝或任何合适数目探针可以是连续的或者分开的固定于芯片表面。最好每种阳性对照探针有多个拷贝、多种不同长度和序列。当固定于芯片上的探针与步骤 b) 制备的目的核苷酸序列或另一核苷酸序列杂交时, 能产生不同量级的由强至弱或由弱至强的杂交信号。

本发明的方法可使用任何适当的阳性对照探针。阳性对照探针最好能与目的核苷酸序列、以目的核苷酸序列为模板扩增产生的核苷酸序列或人工合成的核苷酸序列的一部分互补。



本发明的方法可使用任何适当的阴性对照探针。阴性对照探针最好能与阳性对照探针有 1—3 个碱基对的错配。

本发明的方法可使用任何适当的杂交对照探针。杂交对照探针最好与人工合成的与目的基因无关的核苷酸序列互补，也可与一段人工合成的经标记的序列互补或与一段人工合成的经标记的序列有 1—2 碱基对的错配。

本发明的方法可使用任何适当的固定化对照探针。它是芯片表面化学修饰、点样和固定化过程的内部对照。它不与任何核酸序列杂交。固定化对照探针可一端进行化学修饰，另一端带有可检测的标记。

本发明的方法所用到的芯片可包含任何一条、部分或全部阳性对照探针、阴性对照探针、杂交对照探针、固定化对照探针。例如：一种芯片可包含阳性对照探针、阴性对照探针、杂交对照探针和固定化对照探针。这些对照探针可用任何排布方式固定于芯片上。例如：它们可以被排布在芯片的四角、中心等进行有规律的排布或随机排布。

所述步骤 (d) 中的杂交可以在任何杂交液中进行。如含有 SSC 和表面活性剂的杂交液。杂交液可以含有任何浓度的 SSC (包括氯化钠和柠檬酸钠)，例如  $3 \times -10 \times$  的 SSC。也可以使用任何适当的表面活性剂，如 SDS, Triton100, SLS (十二烷基肌氨酸钠) 等。杂交液中也可以有任何浓度的表面活性剂，如 0.05%—5% (w/w)。

所述步骤 (d) 中的杂交可在任何适当的温度下进行，如  $42^{\circ}\text{C} - 70^{\circ}\text{C}$ 。

本发明的方法还包括杂交后的洗脱步骤。可使用任何适当的洗脱液。洗脱液可含有 0%—2% (w/w) 的表面活性剂，洗脱可持续适当的时间，如 5—30 分钟。

可使用任何适当的方法检测不同探针的固定效率。例如：固定效率可通过分析固定化对照探针的信号来检测。固定化对照探针可带有可检测的标记，如荧光分子。

总体杂交效率，包括与目的核苷酸序列互补的寡核苷酸探针和各种对照探针的杂交效率，可通过任何适当的方法进行检测。例如：可通过分析杂交对照探针和人工合成的与目的基因无同源性的核苷酸序列的杂交信号的进行判断。

总体杂交特异性，包括与目的核苷酸序列互补的寡核苷酸探针和各种对照探针的杂交特异性，可通过任何适当的方法检测。例如：总体杂交特异性可以通过分析阳性对照探针的杂交信号和阴性对照探针的杂交信号的比值，以及阳性杂交对照探针的杂交信号和阴性杂交对照探针的杂交信号的比值来检测。比值提高意味着杂交特异性的提高。

阳性信号可以用任何适当的标准来确定。例如，当一组紧密相关的探针进行杂交时，阳性信号可以以下的标准进行判定：(a) 杂交信号和背景噪音的比值大于 3；(b)



杂交信号相对一个相关的阳性对照探针杂交信号的比值在一个预先确定的范围；(c) 比较所有根据步骤 b 和 c 表现出阳性的探针的杂交信号；或者根据步骤 b 和 c，仅一个探针表现出阳性的时候，比较表现出最强杂交信号的两个探针的杂交信号来确定信号是阳性还是阴性；(d) 在紧密相关的探针组里只有 2 个或少于 2 个阳性信号。

紧密相关的探针组可以根据任何适当标准确定。例如一组设计用来检测特定位点的多态性的探针可被用作一组紧密相关的探针。这种多态性差异包括单碱基变化或者多碱基的变化。正常情况下，碱基对的变化位于探针的长度之内，例如 20 碱基对。

上述预先确定的范围会根据探针的不同而不同。这个范围可以通过经验获得。例如，这个范围可以通过用已知的 HLA 标准序列做多次（例如成百次）杂交试验来获得。

本发明的方法适用于任何基因的分型分析。例如，可以用和目的 HLA 基因互补的寡核苷酸探针来进行 HLA 分型。寡核苷酸探针最好具有以下的序列：(a) 能在高严谨杂交条件下和目的 HLA 序列或其互补序列进行杂交的序列，如序列表中的序列；或 (b) 能够和目的 HLA 序列或其互补序列至少有 90% 的匹配率。寡核苷酸探针最好包括序列表中列出的核酸序列或者其互补序列。芯片可以包括部分或全部序列表中列出的核酸序列或其互补序列。

序列表中，SEQ ID No: 1 至 SEQ ID No: 94 为 HLA-A 的分型探针序列；SEQ ID No: 95 至 SEQ ID No: 170 为 HLA-B 的分型探针序列；SEQ ID No: 171 至 SEQ ID No: 214 为 HLA-DRB1 的分型探针序列。

本发明的另一个目的是提供一种对 HLA 目的基因进行分型的寡核苷酸探针。

本发明所提供的对 HLA 目的基因进行分型的寡核苷酸探针包含如下序列：

a) 能够在高严谨性条件下与目的 HLA 基因核苷酸序列或其互补链进行杂交的序列；所述序列选自序列表中的序列 1—214；

b) 对目的 HLA 核苷酸序列至少能够达到 90% 分辨率的序列，所述目的 HLA 核苷酸序列是正义链或负义链；所述序列选自序列表中的序列 1—214。

所述寡核苷酸探针可以是 DNA、RNA、PNA 或其衍生物。所述寡核苷酸探针最好包括序列表中的序列或其互补序列。探针可以是被标记的，其标记方式包括化学标记、酶标记、免疫标记、放射性标记、荧光标记、发光标记或荧光共振能量转移 (FRET) 标记等。

所述寡核苷酸探针可以用任何方法进行合成。例如：可以化学合成（见 Ausubel 编著的“分子生物学使用操作方法”中的 2.11. 寡核苷酸的合成和纯化，John Wiley & Sons, 公司，2000）、从天然物质中分离、由天然产物重组或综合几种方法合成。人工合成的寡核苷酸也能用 Matteucci 的方法制备（Matteucci 等，J. Am. Chem. Soc.,

3:3185-3191, 1981)。使用自动合成仪也是一种好方法,例如使用ABI公司的DNA合成仪。但是,使用化学方法制备的探针更为合适。

本发明中用于合成所述寡核苷酸探针的碱基是天然的腺嘌呤、鸟嘌呤、胞嘧啶、尿嘧啶、胸腺嘧啶,也可以选择非天然形成的或合成的如:8-氧-鸟嘌呤,6-巯基鸟嘌呤,4-乙酰基胞嘧啶,5-(羧基羟乙基)尿嘧啶,2'-氧-甲基胞嘧啶,5-羧基甲氨基-甲基-2-胸腺嘧啶,5-羧基甲氨基尿嘧啶,二氢尿嘧啶,2-氧-甲基假尿嘧啶, $\beta$ -D-半乳糖基鸟嘌呤,2'-氧-甲基鸟嘌呤,次黄嘌呤, $N_6$ -异戊烯基腺嘌呤,1-甲基腺嘌呤,1-甲基假尿嘧啶,1-甲基鸟嘌呤,1-甲基次黄嘌呤,2,2-二甲基鸟嘌呤,2-甲基腺嘌呤,2-甲基鸟嘌呤,3-甲基胞嘧啶,5-甲基胞嘧啶, $N_6$ -甲基腺嘌呤,7-甲基鸟嘌呤,5-甲基氨基甲基尿嘧啶,5-甲基氨基甲基-2-硫-尿嘧啶, $\beta$ -D-甘露糖基鸟嘌呤,5-甲氧基羧基甲基尿嘧啶,5-甲氧基尿嘧啶,2-甲硫基- $N_6$ -异戊烯基腺嘌呤, $N$ -((9- $\beta$ -D-脱氧呋喃基-2-甲硫基尿嘧啶-6-氨基甲酰基)苏氨酸, $N$ -((9- $\beta$ -D-脱氧呋喃基-2-甲硫基尿嘧啶-6- $N$ -甲基氨基甲酰基)苏氨酸,尿嘧啶-5-氧-醋酸甲基酯,尿嘧啶-5-氧-醋酸,假尿嘧啶,2-硫胞嘧啶,5-甲基-2-硫尿嘧啶,2-硫尿嘧啶,5-甲氧基尿嘧啶,2'-氧-甲基-5-甲基尿嘧啶,2'-氧-甲基尿嘧啶,和3-(3-氨基-3-羧基异丙基)尿嘧啶等。

同样的,寡核苷酸的化学类似物(如磷酸骨架经修饰的核苷)也可用于合成所述寡核苷酸探针。可采取的防止降解的方法有“3'末端帽子法”,参见(Shaw等, *Nucleic Acids Res.*, 19:747, 1991)。有许多新的化学途径可用于将新的连接方式代替磷酸骨架。磷酸骨架的类似物包括磷酸化硫盐,磷酸化二硫盐,甲基磷酸盐,氨基磷酸酯,硼酸磷酸酯,磷酸三酯,醋酸甲酯,3'-硫醋酸甲酯,5'-硫醋酸甲酯,5'-硫醚,碳酸酯,5'- $N$ -氨基甲酸酯,硫酸酯,磺酸酯,氨基磺酸酯,磺胺药,砒,亚硫酸酯,亚砷,硫醚,羟基胺,亚甲基(甲基亚氨基)(MMI) or 亚甲基氧基(甲基亚氨基)(MOMI)连接等。所述寡核苷酸可以是肽核酸,如Milligan等已经描述的(*J. Med. Chem.*, 36:1923, 1993)。所述寡核苷酸探针必须含有一段能与目的DNA分子的一部分进行互补结合的序列。

所述寡核苷酸探针可以是任何长度,没有最长或最短的限制。只要能完成探针的功能,与目的HLA基因杂交即可。所述寡核苷酸探针的长度可短至50, 40, 30, 20, 15或10个核苷酸或更短。同样的,所述寡核苷酸探针也可以有20, 40, 50, 60, 75, 100, 200个核苷酸或更长,例如与整个HLA目的基因一样长。通常,所述寡核苷酸探针至少有14nt,更好的是在18nt以上,最好是有20至30nt,与靶基因互补,没有发卡结构。有的情况下,所述寡核苷酸探针可能最少有30nt或50nt。如果要完全配

对,例如,如果靶基因与探针特异性结合,那么这个杂交双链在平稳的严谨性杂交条件下或探针短至 10—30nt 时能保持相对的稳定性。如果所述寡核苷酸探针可能含有一定程度的错配碱基,比如针对一个可变区域或一个特定属中的所有种的一组序列,则所述寡核苷酸探针可以加大长度(例如 15~40nt)来抵消错配的不利影响。所述寡核苷酸探针不必跨越整个 HLA 基因。任何潜在的可用于特异性地区分 HLA 靶基因的区域都可以用于设计所述寡核苷酸探针。所以,所述寡核苷酸探针能与最少 8nt 的目的区域杂交。而且,所述寡核苷酸探针的片断也可以用于 HLA 基因的分型,只要它们的特异性足够。在低严谨的条件下,所述寡核苷酸探针至少应有 8nt 与 HLA 靶基因杂交,杂交最好在中等严谨或高严谨条件下进行。

本发明的第三个目的是提供一种对 HLA 目的基因进行分型的阵列。

本发明所提供的对 HLA 目的基因进行分型的寡核苷酸探针阵列,包括可以固定大量寡核苷酸探针并适于杂交反应的固相载体,所述探针中至少一条探针具有如下序列:

- a) 能够在高严谨杂交条件下与目的 HLA 核苷酸序列或其互补链进行杂交的序列,如列表中的序列;
- b) 能够与目的 HLA 核苷酸序列或其互补序列至少有 90%匹配的序列,所述序列为列表中的序列。

所述探针可以是 DNA、RNA、PNA 或其衍生物。至少有一条或几条所述探针含有序列表中的核苷酸序列或其互补序列,或至少有一段如列表所示的序列或其互补序列。所述阵列包含序列表中所有序列或其互补序列的探针。

至少有一条或几条或全部所述探针是经过标记的,标记方式包括化学标记、酶标记、免疫标记、放射性标记、荧光标记、发标记,或荧光共振能量转移(FRET)标记。任何适当的支持物如硅基、塑料、玻璃、陶瓷、橡胶和高分子材料等都可以采用。

本发明的方法、探针和探针阵列可以在溶液中使用。更适用于以芯片形式使用,如将探针固定于固体支持物上。

探针可固定于任何适当的物质表面。特别是固体支持物如硅、塑料、玻璃、陶瓷、橡胶和高分子材料,还可以固定于三维多孔的琼脂胶板上,如 Packard 公司的水溶胶芯片(Broude et al., Nucleic Acids Res., 29(19):E92, 2001)。对基于点阵的检测,探针最好固定于固体支持物上构成“生物芯片”。固体支持物可以是生物的、非生物的、有机的、无机的或其任意的组合,存在形式可以是:颗粒状、链状、沉淀物、胶状、薄片、管状、球形、毛细管状、块状、片状、胶片状、板状或玻片状等。

含有探针组的微阵列生物芯片可由许多常规方法制备,如光导法(VLSIP<sup>SM</sup>在



U.S. Patent Nos. 5143854; 5384261 or 5561071 中所描述的)、磁珠法 (U.S. Patent No. 5541061)、点样法 (U.S. Patent No. 5288514; U.S. Patent No. 5556752) 等。

流体通道方法, 如美国专利 No 5677195 和 5384261 所提及的, 可以用来制备具有各种不同探针的微阵列芯片。当探针通过流体通道传送到固体支持物时, 底物的活性区域自动从其它区域分离出来。流体通道方法的详细描述可以在美国专利 NO. 5556752 中找到, 包括用于提高液体通过已设计好的流动路径的保护性遮蔽设备的使用。

点样方法 (Spotting methods) 可被用于制备表面固定有各种探针的微阵列芯片。在这种方法中, 相对小量的反应物被直接传递到支持物上被选择的区域。在某些情况下, 支持物表面可用一种特殊的溶液进行喷射或覆盖。在特定情况下, 一个喷头从一个区域移到另一个区域的过程中会在每一个停顿时喷射必需量的探针或其它试剂。典型的自动装置包括微升级吸量管、纳升级吸量管、喷头卡具和点样针等都可以传递探针或其他溶液到支持物上, 并由一套自动机械系统来控制点到支持物上的具体位置。另外, 这个自动装置还包括一系列的试管、多孔板、管线和一排传递装置, 这样各种试剂能够同时被传递到反应区域。点样方法 (Spotting methods) 在美国专利 Nos. 5288514; 5312233 和 6024138 中有描述。在某些情况下, 流体通道和在支持物上的预定区域进行点样的技术结合起来也能用于制备固定有探针的微阵列芯片。

固定探针的固体支持物最好是扁平形的, 但是它也会存在两种不同的表面构造。例如, 固体支持物存在凹凸不平的区域, 探针在这些区域被合成或被点样。在一些具体情况下, 可以选择具有合适的光吸收特性的固体支持物。例如, 支持物可以是聚合的 Langmuir Blodgett 胶卷、玻璃或功能化玻璃、硅、锗、砷化钙、磷化钙、二氧化硅、氮化硅、改良硅, 或一些凝胶或聚合物如聚四氟乙烯、聚亚乙烯基二氟、聚苯乙烯、聚碳酸盐或它们的组合物。

固体支持物表面可包含活性基团, 如包括羧基、氨基、羟基、硫醇或类似基团, 这些基团适合于和寡核苷酸或核酸上的活性基团连接。那些表面透明而且含有 Si-OH 功能基团的材料 (如: 硅表面) 更适合。

探针可以通过化学或物理的方法, 如通过离子键、共价化合或其它已知的力吸附在支持物上。核酸和寡核苷酸的固定可以通过文献, 例如: Dattagupta et al., *Analytical Biochemistry*, 117:85-89, 1989; Saiki et al., *Proc. Natl. Acad. Sci. USA*, 86:6230-6234, 1989; and Gravitt et al., *J. Clin. Micro.*, 36:3020-3027, 1998 等所描述的一些方法来达到。



探针也可以通过“连接分子”固定到支持物上，例如美国专利 No. 5556752 描述的方法，连接分子可以为探针形成双链的部分提供一个自由的空间，这可能有助于杂交反应的进行。典型的连接分子包括 6—50 个原子，而且包括可固定于支持物表面上的部分。探针固定到支持物上可以通过 C—C 键实现，例如，聚三氟氯乙烯表面；或更好的用硅氧烷键（玻璃或二氧化硅作支持物时使用）。硅氧烷键键合可以通过支持物和连接分子的三氯甲基硅烷或三烷氧基甲基硅烷等基团反应完成。氨基烷基硅烷、羟基烷基硅烷、2-羟乙基-氨基丙基三乙氧基硅烷、羟乙基-氨基丙基三乙氧基硅烷或羟丙基三乙氧基硅烷都是很有用的表面吸附基团。

连接分子还包括一个能被固定到固相载体表面的延伸部分或长链。例如，氨基、羟基、巯基、羧基等基团均适合于将连接分子的延伸部分固定到载体表面。连接分子的延伸部分可以是对聚合物合成反应后续条件保持惰性的多种分子中的一些。长链部分一般都是芳香基、乙炔、乙烯、乙二醇（包含 2—14 单体单元的）低聚物、二（元）胺、二酸、氨基酸、肽或其化合物。

在一些具体情况中，连接分子的延伸部分可以是一段多聚核苷酸或整个连接分子是一段多聚核苷酸。连接分子的延伸部分也可以包括聚乙烯乙二醇、聚核苷、亚烷基、聚乙醇、聚酯、聚氨、聚硫酸酯和其组合物。另外，为了在探针的合成中应用连接分子，连接分子在末梢或末端（相对于固体支持物表面）可以有一个保护基团连接到双功能基团上（如：羟基、氨基或羧酸基）。当去保护并连接后，该末端能被共价键合到寡聚物或探针上。

本发明的方法可用于一次用一条探针去分析一个样品。本发明的方法在高通量的形式下进行更为有利。例如，多个样品能用单一探针同时分析或单一样品能用多个探针同时分析。更好的情况是，大量的样品可以用多个探针同时分析。

杂交可以采用该领域所熟知的任何技术来进行。可以改变杂交条件来提高或降低杂交严谨度、杂交的特异性水平和非特异吸附造成的背景水平（例如，改变杂交或洗脱液的盐浓度或温度）。探针和靶核苷序列之间的杂交可以在任何适宜的严谨性（包括高、中、低）下进行。一般是在高严谨性条件下杂交。

探针和靶核酸之间的杂交可以是同源性的，例如：分子信标技术中用的典型条件（Tyagi S. et al., Nature Biotechnology, 14:303—308, 1996; 美国专利 No. 6150097）和杂交保护分析中用的方法（Gen-Probe 公司；美国专利 No. 6004745）。探针和靶核酸之间的杂交可以是非同原性的（参见那些基于不同类型的硝酸纤维素膜和磁珠所采用特定的条件）。

靶多聚核苷酸序列可以通过和寡核苷酸探针在从高到低的杂交和洗脱严谨性条件



下形成稳定的杂合体而被检测到。通过依赖于探针的杂交反应来作为检测的手段的好处是可以保证特异性。如果探针和靶序列之间是完全互补的（例如是 99%，或更高），则应该用高严谨性的条件进行杂交。如果它们之间存在一些错配，例如当用探针去检测一组不可能完全互补的呈多态性变化的靶序列的时候，杂交的严谨性就应降低。但是，应该优化杂交条件使非特异性杂交达到最小或消除。

那些能够影响杂交和用以消除非特异性杂交的条件在一些文章（Molecular Cloning A Laboratory Manual, second edition, J. Sambrook, E. Fritsch, T. Maniatis, Cold Spring Harbor Laboratory Press, 1989）中已经提到过。通常，低盐浓度和高温可以促进杂交的严谨性。例如：一般来说，高严谨性的杂交条件是指在含有大约  $0.1 \times \text{SSC}$  和  $0.1\% \text{SDS}$  的溶液在  $65^\circ\text{C}$  进行杂交和洗脱；中等严谨的杂交条件是在含有大约  $1 \sim 2 \times \text{SSC}$  和  $0.1\% \text{SDS}$  的溶液于  $50^\circ\text{C} \sim 65^\circ\text{C}$  进行杂交和洗脱；低严谨的杂交条件是在  $2 \times \text{SSC}$  溶液中于  $30^\circ\text{C} \sim 50^\circ\text{C}$  进行杂交和洗脱。

一个可选择的杂交和洗脱的方法是先采用低严谨条件进行杂交（ $5 \times \text{SSPE}$ ,  $0.5\% \text{SDS}$ ），再用含有浓度为  $3\text{M}$  的四甲基氯化铵（ $\text{TMAC}$ ）的溶液来严格洗脱。 $\text{TMAC}$  的作用是消除 A-T 和 G-C 碱基对间结合力的不同，以便在给定的温度下杂交的结果主要与多聚核苷的长度相对应。用  $\text{TMAC}$  可以改变洗脱的温度以达到文献（Wood et al., Proc. Natl. Acad. Sci. USA. 82:1585-1588, 1985）中所期望的严谨性水平。

杂交液可以含有  $25\% \text{甲酰胺}$ 、 $5 \times \text{SSC}$ 、 $5 \times \text{Denhardt}$  溶液、 $100 \mu\text{g/ml}$  单链 DNA、 $5\% \text{右旋糖苷硫酸盐}$ ，或者其它一些已知的对杂交有利的溶液。

探针和靶 HLA 核酸之间的杂交可以采用任何已知的方法进行检测，例如：标记探针、二级探针、靶核酸或者其组合物。同时，在缺少可检测的标记时，也可以用质谱学方法来检测杂交结果（例如，美国专利 No. 6300076 中所描述的）。

可检测的标记是一种可直接或间接进行杂交结果检测的方法。可检测的标记有一种可被测量的物理特性（例如荧光或吸光性）或能参与酶反应。如果用直接标记，则靶核苷酸序列或探针是被标记的，那么通过检测杂合体中的标记物就可以评估杂交的结果。如果用间接标记，则二级探针被标记，那么可以通过检测二级探针和初始杂合体间形成的二级杂交物来实现对杂交结果的检测。

适合的标记包括荧光标记、发色团、发光体、放射性同位素标记、电子密标记试剂、FRET（荧光共振能量迁移）、酶和有特殊结合配体的配基。有特殊用途的标记是：酶活性基团，如酶（Wisdom, Clin. Chem., 22:12431976）；酶的底物（英国专利 No. 1548741）；辅酶（美国专利 No. 4230797 和 4238565）和酶抑制剂（美国专利 No. 4134792）；荧光剂（Soini 和 Hemmila, Clin. Chem., 25:353, 1979）；发色团（包



括藻胆蛋白);发光体,如化学发光和生物发光(Gorus 和 Schram, Clin. Chem., 25:512, 1979 和上述杂志的 1531 页);特殊捆绑配体,如蛋白捆绑配体;抗体;放射性同位素,如  $^3\text{H}$ ,  $^{35}\text{S}$ ,  $^{32}\text{P}$ ,  $^{125}\text{I}$  和  $^{14}\text{C}$ 。这些标记根据他们各自的物理特性(如荧光性,发光性和放射性)或他们的反应活性或结合特性(如抗体,酶,底物,辅酶和抑制剂)等被检测。配体标记对固相状态下寡核苷酸探针的捕获是有用的(例如捕获探针)。相似的标记包括生物素(通过与亲和素或链霉亲和素结合而被检测)和酶类,如:山葵过氧化酶或碱性硫酸酶(通过添加酶的底物产生颜色反应而被检测)。

放射性同位素标记的探针或靶核苷酸可以通过放射自显影来检测。荧光标记的探针和靶核苷酸可以用荧光仪来检测。半抗原和配基(如维生素)标记的核酸可以通过向半抗原或结合了标记配体(如抗毒素蛋白)的蛋白添加半抗原或抗体片段来鉴定。

作为另外一种选择,探针或核酸可用通过其它的试剂来对杂交结果进行检测的标记物进行标记。如果标记物是一种酶,被标记的核酸,如 DNA 最终被放置在一个合适的介质中来决定催化作用的程度。例如,用辅酶来标记的核酸可以通过添加一种酶来检测,因为这个标记是这种酶的辅酶和底物。因而,如果这种酶是磷酸化酶,那介质中能够包括磷酸硝基苯,这样就能通过观察颜色来监控硝基苯酚的产生。如果这种酶是  $\beta$ -半乳糖苷酶,那介质中可以含有硝基苯基磷酸盐,它能够释放硝基苯酚。典型的例子中包括但不限于  $\beta$ -半乳糖苷酶、碱性磷酸酶、木瓜蛋白酶和过氧化物酶。通过原位杂交的研究,表明底物的最终产物是不溶于水的。其它的标记物,如染料等也是可以用作标记物。

标记物可以直接被连接到 DNA 的粘性配合体上,如吖啶染料、菲啶、吩噻、香豆素、吩噻嗪和喹啉可通过共价键等化学键直接联接,或通过与联接在粘性配合体上的微囊体或脂质体结合而达到非直接联接。标记物通过插入化合物连接到 DNA 的粘性配合体上的方法是常用的。具有代表性的插入化合物包括:单或二叠氮羰基甲基物质或乙基物质、单/二叠氮羰基乙基物质、氮羰基乙基物质聚合物(Mitchell et al., J. Am. Chem. Soc., 104:4256, 1982)、4-叠氮羰基-7-氯喹啉、2-叠氮羰基苄、4-氮甲基-4, 5-二甲基当归根素、4-氮甲基-三氧盐(4-氮甲基-4, 5, 8-三甲基补骨脂素)、3-羧基-5 或 8-氨基或羟基补骨脂素。一种特殊的核酸结合性叠氮羰基物质已经被 Forster 等(Nucleic Acid Res., 13:745, 1985)描述过。其它有用的具有光反应活性的插入物质是香豆素,它和嘧啶残留物形成(2+2)环状物。烷化剂,如二氯乙胺和环氧化物或吖啶类,如黄曲霉毒素、多轮烃的环氧化物、丝裂霉素和 norphillin A,也能被用作 DNA 的粘性配合基。特别有用的具有光反应活性的插入物质的是叠氮羰基类物质。它们的活性氮烯结构在长波长的紫外线或可见光下很容易产生,并且



芳基叠氮化物的氮烯结构在它们的重排产物上更易于发生插入反应 (White et al., Meth. Enzymol., 46:644, 1977)。

探针也可以被修饰成特殊的形式, 如: 添加 10—100 个 T 残基用于反相点样, 或与 BSA 结合或固定到磁珠上。

当用间接的方法来检测杂交时, 可被检测的带标记的二级探针可以在探针与靶序列进行最初的杂交之后或杂交过程中加入。在加入二级探针后, 杂交条件可以被改变。杂交之后, 如果初级探针被固定于固体支持物上, 通过洗脱, 没有杂交上的二级探针能够与初级探针分离开来。对固体支持物上的特定区域的标记程度的检测就可以显示样品中靶核苷酸序列和探针的杂交情况。

可被检测的带标记的二级探针是一种特殊的探针。相对来说, 可检测的标记探针可以是退化探针, 如美国专利 No. 5348855 中描述的全基因组 DNA 的混合物。近来的研究发现, 如果二级探针含有双股 DNA, 可以通过插入染料来标记。如上面提到的, 首选的 DNA 结合配体就是插入化合物。

二级探针也可以是一个随机核苷酸探针序列库。二级探针的长度应视一级探针的长度和组成或固体支持物上由二级探针鉴定的靶核苷酸序列而定。这样的探针库的探针最好 3' 或 5' 末端的一端用光反应试剂标记, 另一端用可检测试剂如荧光剂、酶、染料、发光体或其它已知的可检测物进行标记。

用作标记核酸的特殊序列是可变的。例如氨基取代的补骨质素能首先被核酸光化学偶联, 其产物带有能结合标记物的氨基基团等, 通过光化学反应, DNA 结合配体标记到核酸上, 即补骨质素首先和标记物如酶结合, 然后和核酸结合。其好处在于, DNA 结合配体首先和标记物结合, 然后再和核酸探针结合。例如, 由于生物素带有羧基基团, 它和香豆素通过酰胺或酯的形式结合, 而不影响香豆素的光反应活性或生物素的生物学活性。氮甲基当归素、补骨质素和菲啶衍生物可同样结合到标记物上, 如菲啶氯化物和氮丙基甲基氯化物等衍生物 (Hertzberg et al., J. Amer. Chem. Soc., 104:313, 1982)。另外, 双功能试剂如二硫代双丁二酰胺基丙酸或 1,4-丁二醇二环氧甘油醚可在适当的溶剂、反应物配比和反应条件下直接把 DNA 结合配体和有氨基残基的标记物连接。某些双功能试剂如戊二醛可能不合适, 因为它们发生耦合反应时可能对核酸发生修饰作用而影响分析结果。常规的预防措施能防止这样的情况发生。

同样有利的是, DNA 结合配体也能够通过连接分子与标记物相连, 此连接分子的长度可达 40 个原子, 但 2 到 20 个原子的长度更合适。组成连接分子的原子包括但不限于碳、氧、氮、硫。此连接分子可以是多功能基团, 例如肽、碳氢化合物、多元醇、多元醚、多胺、碳水化合物类如甘氨酸聚合物或其它寡肽、羧基肽、 $\alpha$ -氨基-碳



酰基或类似的物质。糖类、聚乙烯氧化基团、甘油基、季戊四醇基团等也可以作为连接分子。连接分子能直接连到核酸分子或/和标记物上。连接分子可包括双功能基偶联试剂，如：二硫代双丁二酰胺基丙酸或 1,4-丁二醇二环氧甘油醚，二异氰酸酯、碳二酰亚胺、乙二醛、戊二醛之类。

不能直接进行检测的二级探针因能量转移而可被检测，如 Tyagi 和 Kramer 描述的分子信标法 (Nature Biotech., 14:303-309, 1996 or U.S. Patent No. 5119801 and 5312728)。上述任何荧光共振能量转移 (FRET) 检测系统都可以应用于本方法。例如可以使用 AlphaScreen™ 系统。AlphaScreen 是一种近距离放大发光均相检测 (Amplified Luminescent Proximity Homogeneous Assay) 方法。用 680nm 激光照射时，供体小珠上的光敏剂能将周围的氧变为氧单态。激发态氧分子在衰变前发出 250nm (接近小珠直径) 的光。如果受体小珠与供体小珠很接近，氧单态分子就能与受体上的化学发光基团作用，将能量很快转移到同一小珠的荧光物质上。荧光物质将光波转换为 520—620nm。整个过程的半衰期是 0.3 秒。荧光共振能量转移 (FRET) 供体/受体对可以是荧光素 (供体) 和四甲基罗丹明 (受体)，其有效波长距离是 55 埃；IAEDANS (供体) 和荧光素 (受体)，其有效波长距离是 46 埃；荧光素 (供体) 和 QSY-7 染料 (受体)，其有效波长距离是 61 埃。

本发明也可用于核酸的定量检测。结合于微阵列点上的二级探针的数量能够被检测，并且与目的核酸分子的数量相关。根据含有已知数量的目的核酸的对照的数量可以对样品进行稀释。其精确程度与操作者的熟练程度有关。微阵列分析中，将探针阵列放在 X 光胶片或显像仪旁边可显示出标记物，从而确定固定探针的位置。荧光能用光电倍增管 CCD 或激光扫描探测。

任何合适的样品，包括来自于人类、动物、环境产物 (如泥土或水) 的样品，都能用本发明的方法分析。测试样品包括各种体液，如：尿液、血液、精液、脑脊髓液、脓、羊水、眼泪或半固体或液体排出物如痰、唾液、呼吸产物、阴道尿道排出物；粪便或固体组织样品，如：活体组织或绒毛膜样品。检测样品还包括皮肤、生殖器或咽喉刮取物。

可使用多种方法提取被检样品的核酸，如 (Ausubel (Ed.) Current Protocols in Molecular Biology, 2. Preparation and Analysis of DNA and 4. Preparation and Analysis of RNA, John Wiley & Sons, Inc. (2000))。所得到的目的核酸可以是经扩增或纯化的 DNA 或 RNA。

上述方法得到的核酸可用色谱或其它仪器测量。可以使用各种扩增反应得到扩增产物，如 PCR (Polymerase Chain Reaction, U.S. Patent Nos. 4,683,195, 4,683,202,



4,800,159 and 4,965,188), NASBA (Nucleic Acid Sequence Based Amplification, U.S. Patent No. 5,130,238), TMA (Transcription Mediated Amplification) (Kwoh et al., Proc. Natl. Acad. Sci., USA, 86:1173-1177 (1989)), SDA (Strand Displacement Amplification, described by Walker et al., U.S. Patent No. 5,270,184), tSDA (thermophilic Strand Displacement Amplification (U.S. Patent No. 5,648,211 and Euro. Patent No. EP 0 684315), SSSR (Self-Sustained Sequence Replication) (U. S. Patent No. 6,156,508)。

本发明的探针可以包装为试剂盒的形式,最好包括如何使用这些探针去检测目的基因的说明书。试剂盒的各成分可以装在一个共同的箱子里,试剂盒可选择性包括本发明描述的检测方法,如二级探针,和/或标记试剂和进行标记的操作方法(如放射性标记物、酶反应底物、抗体等这类物质)。

本发明的分型方法可用于建设人类骨髓干细胞库和脐带血干细胞库、器官移植检测、骨髓移植检测、自身免疫疾病研究、病毒感染,以及癌症研究、疾病易感性研究、法医鉴定、亲子鉴定和人类基因组研究等方面。

#### 附图说明

图1为纳米磁珠捕获白细胞的照片

图2为几种不同纳米磁珠捕获白细胞后直接进行不对称PCR的检测结果

图3为Fe<sub>3</sub>O<sub>4</sub>磁珠捕获白细胞后直接进行不对称PCR的检测结果。

图4为点有144条探针的芯片的杂交结果

图5A为DHPLC对PBH\_0303019探针的检测结果

图5B为DHPLC对PBH\_0301119探针的检测结果

图6A为纯度很高的探针

图6B为含少量杂质的探针

图6C为含较多杂质的探针

图6D为纯度很差的探针

#### 具体实施方式

##### 定义

除非特别指出,这里所有用到的技术和科学术语的含义与熟悉本领域知识的人所理解的一致。如果本专利申请书中所列出的某个定义与上述参考文献中所提到的一个定义在意思上相反或者不一致,那么将以本说明书中的定义为准。

“一个”,意味着“至少一个”或者“一个或更多”。

“引物”是一种寡核苷酸,它能够与靶序列结合,它可在扩增反应中作为核酸扩



增的引物。

“探针”是指一种能够与靶序列或目的序列杂交的寡核苷酸。探针不像引物那样在聚合酶的作用下去延伸、扩增靶序列。但是，对于熟知本领域的人都很清楚，在许多情况下探针和引物在结构上是相似的或完全相同。

“靶序列或目的序列”，是指能够和探针特异性的结合一种核酸序列。

“阳性对照探针”是一种能够与一组（或一类）靶序列上的保守序列杂交的探针。

“阴性对照探针”是指和阳性探针相比，在序列上有一个或多个碱基改变的探针。与阳性对照探针只有一个碱基不同的阴性探针要更合适一些。再比如阴性对照探针可以是与原靶序列无同源性的序列的同源序列。举一个特殊的例子，当两种阳性对照探针：如一种强阳、一种弱阳，及一种单碱基不同的阴性对照探针被一起使用时，强阳对照探针和阴性对照探针在杂交过程中用于评估杂交效率，弱阳性对照探针则用来评估待检测探针的杂交信号，例如非对照的探针杂交信号将会以此来评估。待检测探针的杂交信号和弱阳对照探针的杂交信号的比值在一定数值范围内用于评估待检测探针的杂交强度。

“杂交对照探针”是指用于评估待检测探针和靶序列的杂交效率的探针。例如，如果靶序列是一种 HLA 序列，则杂交对照探针的序列最好是与任何 HLA 序列都不同源的序列。杂交对照探针可以被氨基修饰，然后被用于（或固定于）芯片表面，其点样浓度和/或点样过程与其它探针相同或相似，其中也包括待检测探针。另一种标记探针，如 HEX 标记探针，它是一种标准的杂交对照探针，它被以一种浓度或比例添加到杂交液中，而这个浓度或比例与其它探针的浓度或比例相一致。这样，整个的杂交过程可以被检测到。杂交对照探针也被用于指导或决定探针在芯片上的位置。

“固定化质控探针”是指用于评估芯片固定化过程的探针，它不参与任何的杂交过程。固定化质控探针的一端被修饰，如氨基修饰，以利于其在芯片表面的固定，而它的另一端则被一种易被检测的标记物标记，如 Hex。

“Hex”指六氯荧光素，是一种常用的荧光标记物。

“互补”是指两种核酸序列至少有 50% 的一致性，一般二者至少有 60%、70%、80%、90%、95%、96%、97%、98% 或 100% 的一致性。互补也指两种核苷酸序列能够在低、中和/或高三种严格程度的条件下杂交。

“充分匹配”是指两种核苷酸序列有至少 90% 的统一性，即“充分匹配”是指两种核苷酸序列可以在高严谨的杂交条件下杂交。

“两种完全匹配的核酸序列”是指两条核酸单链根据 Watson-Crick 碱基匹配原则而生成的核酸双链，如在 DNA:DNA 中是 A-T 和 C-G 的相互配对，在 DNA:RNA 或



RNA:RNA 中则是 A-U 和 C-G 的相互配对, 除此之外, 不会有其他任何的碱基配对情况。

“杂交严格性”是指如下条件:

- 1) 高严谨性:  $0.1 \times$  SSPE, 0.1% SDS, 65°C;
- 2) 中严谨性:  $0.2 \times$  SSPE, 0.1% SDS, 50°C (也被称为温和的严谨性);
- 3) 低严谨性:  $1.0 \times$  SSPE, 0.1% SDS, 50°C。

通过选择缓冲液、盐浓度和温度来获得相同的要求标准。(Ausubel (Ed.) Current Protocols in Molecular Biology, 2.9A. Southern Blotting, 2.9B. Dot and Slot Blotting of DNA and 2.10. Hybridization Analysis of DNA Blots, John Wiley & Sons, Inc. 2000)。

“退火温度”是指核酸双链, 如 DNA:DNA, DNA:RNA 和 RNA:RNA, 在变性过程中温度变化范围的中值时的温度。这里指的探针的  $T_m$  指的是发生杂交反应的探针的  $T_m$ 。

“评价”是指对探针和靶核酸序列之间形成的杂合体的定量和/或定性的检测。例如检测一个杂合体的数量或浓度的绝对值; 检测某一个指数, 比值, 百分数, 图表或其它的数值来显示杂交的效果。评价可以是直接的或间接的, 实际上化学类物质的检测可以不必检测杂交体本身, 可以通过衍生化、还原化、或探针和/或靶核苷酸序列或其它的物质的减少或消失等来评价。

“磁性物质”是指那些具有磁铁的特性、与磁铁或磁性有关、能够用磁力进行诱导或操纵的任何物质。

“可磁化的物质”是指当悬浮或自由放置在磁场内能被磁化和产生磁距的任何物质。这些磁化物质包括顺磁物质, 磁铁和亚铁磁物质, 但不仅仅限于这些物质。

“顺磁物质”指其单个原子、离子或分子具有永久的偶磁极的物质, 在外加磁场的的作用下, 这些偶磁极随机取向, 而物质本身整体上不表现出定向的磁场。这种随机取向是由物质内部热运动引起的。外加磁场后, 原子偶磁极都在磁作用下平行于外磁场, 因为这种状态下能量最低, 这也增加了磁场的磁化系数。关于顺磁物质和顺磁性的详细介绍, 能在许多文献中找到, 例如: 由 B. I Bleaney 和 B. Bleaney 编著的“电学和磁学”第六章, P169~171, 1975, 牛津出版社出版。

“铁磁性物质”指具有显著正磁化系数的物质。磁系数由磁场强度决定。另外, 铁磁性物质能在外磁场消失后保持磁性, 更多细节详见 B. I Bleaney 和 B. Bleaney 编著的“电学和磁学”第六章, P171~174, 1975, 牛津出版社出版。

“亚铁磁物质”指具有自发的永久磁性, 与普通磁铁特性相同的物质。但这些自发的磁性并不一定是由于该物质内部偶磁极平行排列而产生的。关于铁磁物质和亚铁磁物质更详尽的资料可见于许多文献中, 如: B. I Bleaney 和 B. Bleaney 编著的“电

学和磁学”第十六章, P519~524, 1975, 牛津出版社出版。

“金属氧化物颗粒”指以颗粒形式存在的某种金属的任何氧化形式。某些特殊的金属氧化物颗粒具有顺磁性或超顺磁性。顺磁颗粒指易被磁场磁化但自身不能维持一个磁性区域的颗粒。顺磁颗粒是由顺磁物质组成的颗粒。顺磁颗粒包括金属氧化物, 如  $\text{Fe}_3\text{O}_4$  颗粒, 和合金颗粒如  $\text{CoTaZr}$  合金颗粒等。

“样品, 如新鲜全血”是指从其天然来源处分离、获得后 12 小时以内的样品, 当然也包括 10、5、4、3、2、1 小时, 30、20、10、5、2 或者 1 分钟以内的样品。

“低温保存”指将样品保存在  $0^\circ\text{C}$  或更低的温度。

### 实施例 1、HLA 基因分型

#### (1) 分离人全血中的白细胞及目的基因的扩增

本实施例中白细胞用一种  $\text{Fe}_3\text{O}_4$  磁珠或  $\text{Fe}_3\text{O}_4$  磁珠、醛基包覆的  $\text{Fe}_3\text{O}_4$  磁珠、羧基包覆的  $\text{Fe}_3\text{O}_4$  磁珠、环氧基包覆的  $\text{Fe}_3\text{O}_4$  磁珠、羧基包覆的  $\text{Fe}_3\text{O}_4$  磁珠等种磁珠进行分离 (例如图 2 所用的三种磁珠分别是: I  $\text{Fe}_3\text{O}_4$  裸珠, II 醛基包覆的  $\text{Fe}_3\text{O}_4$  磁珠, III 羧基包覆的  $\text{Fe}_3\text{O}_4$  磁珠)。

5  $\mu\text{l}$  用 ACD (23 mM 柠檬酸, 80 mM 葡萄糖, 45 mM 柠檬酸钠) 抗凝的全血与 20  $\mu\text{l}$  磁珠 (悬于 pH6.0 的 TE 缓冲液制成 15 mg/ml 悬液) 在一支 1.5ml 的 EP 管中于振荡器上轻微混匀 10 秒钟, 静置 3 分钟。然后用磁力架吸附住磁珠, 弃掉上清。用 PBS 洗涤磁珠两次。然后将磁珠重新混悬于 100  $\mu\text{l}$  TE 缓冲液中, 取 5  $\mu\text{l}$  这种混悬液作为 PCR 的模板。用 25  $\mu\text{l}$  体系进行 PCR 扩增, PCR 混合物组成是: 2.5  $\mu\text{l}$   $10\times\text{PCR}$  缓冲液, 0.5  $\mu\text{l}$  10 mM dNTP, 0.5  $\mu\text{l}$  1 mM Cy5-dCTP, 0.5  $\mu\text{l}$  100 ng/ $\mu\text{l}$  模板 DNA, 0.5  $\mu\text{l}$  1  $\mu\text{M}$  上游引物 (5'-TCC CCA GAC GCC GAG GAT GGC C-3'), 2.5  $\mu\text{l}$  10  $\mu\text{M}$  下游引物 (5'-CCC GTG GCC CCT GGT ACC CG-3'), 0.5  $\mu\text{l}$  5 U/ $\mu\text{l}$  Taq 酶, 加无菌双蒸水至 25  $\mu\text{l}$ 。PCR 扩增程序是:  $96^\circ\text{C}$  变性 3min; 25 个循环,  $96^\circ\text{C}$  变性 25 秒,  $71^\circ\text{C}$  退火 45 秒,  $72^\circ\text{C}$  延伸 30 秒; 9 个循环,  $96^\circ\text{C}$  变性 25 秒,  $68^\circ\text{C}$  退火 60 秒,  $72^\circ\text{C}$  延伸 120 秒;  $72^\circ\text{C}$  延伸 8 分钟。

PCR 结束后用 1.2% 琼脂糖凝胶检测扩增结果, PCR 产物的检测结果如图 2 和图 3 所示, 表明纳米磁珠捕获的白细胞可以直接作为模板进行不对称 PCR 反应。图 2 中显示的是三种  $\text{Fe}_3\text{O}_4$  磁珠分别在三种不同模板量情况下的扩增结果。图 3 显示的是一种  $\text{Fe}_3\text{O}_4$  磁珠捕获白细胞后直接扩增的结果, 其中 3、5、7 指不同的模板量。图 2 和图 3 中, M 表示 DL-2000 Marker, I、II、III 代表不同的磁珠, 3、5、7 代表不同的模板量。可以看出图中的 PCR 产物有两条带, 分别位于 1000bp 和 500bp 左右, 这与理论数值一致。这些 PCR 产物可以不经纯化直接用于下面的杂交步骤。

#### (2) 杂交及 HLA 基因分型

5  $\mu\text{l}$  PCR 产物与 3  $\mu\text{l}$  杂交液混合,  $98^\circ\text{C}$  保温 5 分钟, 然后使这些 PCR 产物和芯片



上固定的 144 条探针（探针的种类详见表 3）在 65℃ 杂交 1 小时。杂交结束后，用去离子水洗去未杂交的 PCR 产物，再用芯片洗涤液于 45℃ 洗涤 10 分钟，再用去离子水漂洗一次，干燥玻片。

用特殊的扫描仪（这里用的是 Scanarray 4000 共聚焦激光扫描仪，也可以用其它的激光扫描仪）扫描杂交后的芯片得到杂交图结果如图 4 所示，再用专用软件（如 Genepix3.0）处理图像得到数据文件，分析得到的数据文件就可以得到目的基因的基因型。如图 4 所示的分型结果是 HLA-A\*01/29 的分型结果。

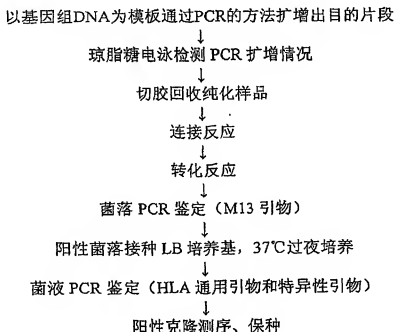
#### 实施例 2、探针质量的检测

PBH\_0303019 和 PBH\_0301119（由上海博亚生物工程公司合成）这两条一端经氨基修饰的寡核苷酸探针溶于水配成 10μM 溶液，用 WAVE® 核酸片段分析系统（Transgenomic, USA）进行分析。柱温 80℃，探针用梯度浓度的乙氧缓冲液洗涤。用紫外监测器在 260nm 进行检测。随洗脱时间不同而呈现出的峰的数目和形状就可以表明探针的质量。PBH\_0303019 和 PBH\_0301119 的质量检测结果如图 5A 和图 5B 所示，表明 DHPLC 不但可以有效地检测芯片所用的探针的纯度，同时可以准确检测出探针所含杂质的具体含量。如图 5A 和 5B 中所示探针的有效含量分别为 93.35% 和 64.8%。

#### 实施例 3、HLA\_A, B, DRB1 三个基因座位的分型探针

表 1 中 PBH\_0301xxx 为 HLA-A 区的分型探针，PBH\_0302xxx 为 HLA-B 区的分型探针，PBH\_0303xxx 为 HLA-DRB 区的分型探针。

#### 实施例 4、克隆标准 HLA 等位基因（具体流程如下）





为了检测 HLA\_A, B, DRB1 三个基因座位的分型探针和控制芯片的质量而构建了一个 HLA 标准等位基因库, 其中 HLA-A 区共克隆成功 43 个等位基因, 分布于 19 组, 可以覆盖我国 HLA 中分辨率分型 95.0% 的型别; HLA\_B 区共克隆成功 47 个等位基因, 分布于 29 组, 可以覆盖中分辨率分型 80% 的型别; HLA\_DR 区共克隆成功 22 个等位基因, 分布于 14 组, 可以覆盖中分辨率分型 87.5% 的型别。根据第十二届国际组织相容性工作委员会 (www.ihwg.org) 提供的人类 HLA 等位基因序列来设计表 1 中的全部探针。然后把这些探针固定在经醛基化修饰的玻璃芯片表面从而构成芯片。

选择能够产生较强阳性信号和最小假阳性信号的杂交条件和洗脱条件。用较高的杂交温度 (如 65℃) 进行杂交, 用合适的温度 (如 45℃) 和较强的离子强度 (如 0.2%SDS, 0.1xSSC) 进行洗涤, 以便保证较好的阳性信号, 并有效降低假阳性信号。

含有靶基因的样品和芯片的杂交、用一种溶液洗涤玻片、信号扫描并通过靶核苷酸序列和固定在芯片上的探针之间的杂交结果来确定靶基因的类型。例如可以取含有靶基因的 PCR 产物 5μl 与 5μl 杂交缓冲液混合后与芯片在 65℃ 杂交 1 小时, 用含 0.2% SDS 和 0.1xSSC 的洗涤溶液在 45℃ 洗涤 10 分钟, 再用 1xSSC 洗涤 5 分钟, 再用水漂洗 1 分钟, 用离心机将玻片甩干, 用芯片扫描仪进行扫描, 用分析软件分析得到的图谱, 将得到的数据输入特定的 HLA 分析软件判断分型的结果。

对这些克隆的标准 HLA 等位基因测序发现: 许多序列和第十二届国际组织委员会提供的序列不匹配。

所设计的 HLA\_A, B, DRB1 三个基因座位的分型探针如表 1。

表 1. HLA\_A, B, DRB1 三个基因座位的分型探针

探针名称	HLA_A probe
PBH_0301001	CCTGCGCTCTTGGACCGC
PBH_0301001a	CCTGCGCTCTTGGACCGCG
PBH_0301001b	CCTCCTGCGCTCTTGGACCG
PBH_0301001c	CCTGCGCTCTTGGACC
PBH_0301001d	CGTGTCCCGGCCCGGC
PBH_0301001e	ATGGAGCCGCGGGCGC
PBH_0301001g	CCT GCG CTC TTG GAC CGC GG
PBH_0301001comp	GCGGTCCAAGAGCGCAGG
PBH_0301002a	CCTGCGCTTTTGGACCGC
PBH_0301002B	CCT GCG CTG TTG GAC CGC
PBH_0301003	GCAGGAGAGGCCTGAGTATTGG



PBH\_0301004  
PBH\_0301004  
PBH\_0301101  
PBH\_0301103  
PBH\_0301105  
PBH\_0301106  
PBH\_0301107  
PBH\_0301111  
PBH\_0301112  
PBH\_0301114  
PBH\_0301115  
PBH\_0301116  
PBH\_0301117  
PBH\_0301119  
PBH\_0301120  
PBH\_0301121  
PBH\_0301122  
PBH\_0301123  
PBH\_0301125  
PBH\_0301127  
PBH\_0301128  
PBH\_0301129  
PBH\_0301130  
PBH\_0301131  
PBH\_0301133  
PBH\_0301136  
PBH\_0301137  
PBH\_0301138  
PBH\_0301139  
PBH\_0301140  
PBH\_0301141  
PBH\_0301142  
PBH\_0301143  
PBH\_0301144  
PBH\_0301145

C ACC ATC AGATAATGTATGGCTGC  
C ACC ATC CAG ATA ATG TAT GGC TGC  
TTCTACACCTCCGTGTCCTCCG  
CGCTTCATCGCAGTGGGCT  
CGAGCCAGAAGATGGAGCC  
CCGCGGGCACCGTGGATA  
GCAGGAGGGTCCGGAGTATT  
GACGTGGGGCCGGACGGG  
GACGGGCGCCTCCTCCGC  
CGGGTACCACAGTACGCCTA  
GGTACCGGCAGGACGCCTA  
CGCCCTGAACGAGGACCTG  
CGGACATGGCAGCTCAGATC  
CCACCAAGCACAAGTGGGA  
AAGTGGGAGACGGCCCATG  
AGGCGGCCCGTGTGGCGG  
AGGCGGTCCATGCGGCGG  
CGGCCCATGAGGCGGAGC  
TACCTGGATGGCAGCTGCG  
CTGGAGGGCGAGTGCCTGG  
TGCGTGGACGGGCTCCGC  
GTATTCTACACCTCCGTGTCCCG  
CGAGCGGTTTGACAGCGAC  
CGTGCGGTTTCGACAGCGAC  
CGTGGGGCCGGACGGG  
AGGCGGTCCATGCGGCG  
CCCGGCCCGGGGAGCCCC  
CCGCGGGCGCCGTGGATA  
TGGGACGAGGAGACAGGGA  
TGGGACCAGGAGACACGGA  
TGGGGACCTGCGCGGCTA  
GACGTGGGGTGGACGGG  
GACGGGCGCTCCTCCGC  
GCGGGTACCAGCAGGACGC  
CGCCCTGAAAGAGGACCTG





PBH\_0301146  
PBH\_0301146a  
PBH\_0301147  
PBH\_0301148  
PBH\_0301148a  
PBH\_0301149  
PBH\_0301149a  
PBH\_0301150  
PBH\_0301150a  
PBH\_0301151  
PBH\_0301151a  
PBH\_0301152  
PBH\_0301153  
PBH\_0301154  
PBH\_0301155  
PBH\_0301155a  
PBH\_0301156  
PBH\_0301156a  
PBH\_0301157  
PBH\_0301171  
PBH\_0301172  
PBH\_0301173  
PBH\_0301174  
PBH\_0301175  
PBH\_0301176  
PBH\_0301177  
PBH\_0301178  
PBH\_0301179  
PBH\_0301180  
PBH\_0301181  
PBH\_0301182  
PBH\_0301183  
PBH\_0301184  
PBH\_0301185  
PBH\_0301186

AGCTCAGATCACCAAGCGCA  
TCAGATCACCAAGCGCAAGAG  
AGCTCAGATCACCGAGCGCA  
GGCTCAGATCACCCAGCGCA  
TCAGATCACCCAGCGCAAGTG  
AGACGGCCCATGAGGCG  
AGACGGCCCATGAGGCG  
GCGGAGCAGCGGAGAGTCT  
AGACGGCCCATGAGGCGG  
GCGGAGCAGTTGAGAGCCT  
GCGGAGCAGTTGAGAGCC  
GCGGAGCAGTGGAGAGCCT  
TACCTGGAGGGCACGTGCG  
TGCGTGGAGTGGCTCCGC  
TCACCGAGTGGACCTGGGG  
CCGAGTGGACCTGGGGACC  
TGACCGAGAGAACCTGCGG  
CCGAGAGAACCTGCCGATCG  
GAAGGCCCACTCACAGACTG  
TATTTCTTCACATCCGTGTCCCG  
TCTACACTTCGGTTTCCCGGC  
CTACACCTCCATGTCCCGGC  
CCGGAACACACGGAAAGTGAA  
ATTGGGACGGGGAGACAG  
GACACGGAATATGAAGGCCCA  
GACACGGAATGTGAAGGCC  
TCACAGACTCACCGAGTGGACC  
TCACAGATTGACCGAGTGGACC  
TCACAGACTGACCGAGTGGACC  
CGAGCGAACCTGGGGACC  
CCGAGAGAGCCTGCGGATC  
ACCGAGAGAACCTGGGGACC  
GTGGACCTGGCGACCTGC  
CACCGTCCAGAGGATGTATGGC  
ACCAGCAGGACGCTTACGACG



PBH_0301187	TCGCCCTTGAACGAGGACCTG
PBH_0301188	CCTGCGCTCTTGGACCCG
PBH_0301189	TCAGACCACCAAGCACAAAGTGG
PBH_0301190	GAGGCGGCCCATGTGGC
PBH_0301191	GGCCCATGCGGCGGAGC
PBH_0301192	GCGGCCCGTCGGGCGGA
PBH_0301193	GCACGTGCGTGGAGTGGC
PBH_0301194	GCCGTGCGTGGACGGGC
PBH_0301195	GGCGAGTGCCTGGAGTGGC
PBH_0301196	GCACGTGCGTGGACGGGC
PBH_0301197	GCCGTGCGTGGAGTGGC
PBH_0301198	GGCGAGTGCCTGGACGGGC
PBH_0301199	AGACACGGAAGTGAAGGCC

**探针名称****HLA\_B probes**

PBH_0302001	TGGCCCTGACCGAGACCTGGGC
PBH_0302001a	CTACAACCGAGCGAGGCCG
PBH_0302002(negative)	GCCCTGACCCAGACCTGGG
PBH_0302003	CCCGAACCCTCCTCTG
PBH_0302004	CCCGAACCGTCTCTG
PBH_0302005	TGCTCTCGGCGGCCCTG
PBH_0302006	TGCTCTCGGAGCCCTGG
PBH_0302007	GGGGGCGAGTGGCCCT
PBH_0302008	TGAGGTATTTCGACACCGCCA
PBH_0302009	TGAGGTATTCTACACCGCCATG
PBH_0302010	TTCCACACCTCCGTGTCCC
PBH_0302011	TCTACACCGCCATGTCCCG
PBH_0302012	TCTACACCTCCGTGTCCCG
PBH_0302013	CCGTTTCATCTCAGTGGGCTAC
PBH_0302014	CGCTTCATCACCGTGGGCT
PBH_0302015	CGCTTCATCGCAGTGGGCT
PBH_0302016	TACGTGGACGGACCCAGTT
PBH_0302017	CGTGGACGACACCCAGTTCG
PBH_0302018	GGACGACACGCTGTTCTGTGA




PBH\_0302019 TGGACGACACGCAAGTTCGTG  
PBH\_0302020 GCGACGCCACGAGTCCG  
PBH\_0302021 GCGACGCCCGGAGTCC  
PBH\_0302022 GAGTCCGAGAGAGGAGCCGC  
PBH\_0302023 CCGAGGAAGGAGCCGCG  
PBH\_0302024 AGGATGGCGCCCCGG  
PBH\_0302025 GGACGGAGCCCCGGGC  
PBH\_0302026 CGGCGCGCGTGGATAGAG  
PBH\_0302027 CGGCGCGCATGGATAGAG  
PBH\_0302028 GGGGCCCGAATATTGGGAC  
PBH\_0302029 GGGGCCCGAGTATTGGGAC  
PBH\_0302030 GGGACCGGAGACACAGATCT  
PBH\_0302031 TGGGACCGGAACACACAGATC  
PBH\_0302032 ACACAGAAGTACAAGCGCCAGG  
PBH\_0302033 ACACGGAACATGAAGGCCTCC  
PBH\_0302034 CACACAGATCTTCAAGACCAACAC  
PBH\_0302035 ATCTGCAAGGCCAAGGCACA  
PBH\_0302036 TACAAGGCCCGAGGCACAGACT  
PBH\_0302037 ACACAGACTGACCGAGAG  
PBH\_0302038 CACACAGACTTACCGAGAGAGCC  
PBH\_0302039 GCACCGCGCTCCGCTA  
PBH\_0302040 CGGACCCTGCTCCGCTACT  
PBH\_0302041 ACCTGCGGATCGCGCTC  
PBH\_0302042 CGGAACCTGCGCGGCT  
PBH\_0302043 CGGGTCTCACATCATCCAGAGG  
PBH\_0302044 GGGTCTCACACCTCCAGAGG  
PBH\_0302045 TCACACTTGGCAGACGATGTATG  
PBH\_0302046 ACACCCTCCAGAGGATGTACGG  
PBH\_0302047 CGACCTGGGGCCCGAC  
PBH\_0302048 CGACGTGGGGCCGGAC  
PBH\_0302049 GGGTACCACCAGGACGCCT  
PBH\_0302050 CGGGTATGACCAGGACGCC  
PBH\_0302051 GGGCATGACCACTCCGCC  
PBH\_0302052 GCGGGTATAACCAGTTCCGCC  
PBH\_0302053 GAGGACCTGCGCTCCTGGA

38

PBH_0302054	GAGGACCTGAGCTCCTGGA
PBH_0302055	GGACCGCCGCGGACAC
PBH_0302056	GGACCGCGGCGGACAC
PBH_0302057	CGGACACGCGGCTCAG
PBH_0302058	CGGACACCGCGGCTCAG
PBH_0302059	GGCCCGTGAGGCGGAG
PBH_0302060	GGCCCGTGTGGCGGAG
PBH_0302061	GCGGAGCAGGACAGAGCCTA
PBH_0302062	GCGGAGCAGTGGAGAGCCTA
PBH_0302063	GCGGAGCAGCTGAGAGCCTA
PBH_0302064	AGCAGCTGAGAACCTACCTGGAG
PBH_0302065	AGCAGCTGAGAGCCTACCTGGAG
PBH_0302066	GGAGGGCGAGTGCGTG
PBH_0302067	GGAGGGCACGTGCGTG
PBH_0302068	GGAGGGCCTGTGCGTG
PBH_0302069	CGTGGAAGTCCGTCGAG
PBH_0302070	CGTGGAAGTGGTCCGTCGAG
PBH_0302071	CTCCGAGACACCTGGAGAAC
PBH_0302072	GCTCCGAGATACCTGGAGAA
PBH_0302073	AGGACAAGCTGGAGCGCG
PBH_0302074	GGACACGCTGGAGCGCG
PBH_0302075	GGAGACGCTGCAGCGCG

探针名称	HLA_DRB1 probes
PBH_0303001	CTTGTTGGCAGCTTAAGTTTGAATGT
PBH_0303002	TGGAGTACTCTACGTCTGAGTGTCA
PBH_0303003	GGAGCAGGTTAAACATGAGTGT
PBH_0303004	CCTGTGGCAGGGTAAGTATAAGT
PBH_0303005	TTGGAGTACTCTACGGGTGAGTG
PBH_0303006	CCTGTGGCAGCCTAAGAGGG
PBH_0303007	CCTGGAGCAGGCGCGG
PBH_0303008	CCTGGAAGACGAGCGGGC
PBH_0303009	CCAGGAGGAGAACGTGCGC
PBH_0303010	CCTGGAAGACAGGCGGGC
PBH_0303011	CGGTGTCTGGAAGATGCATC


  
 PBH\_0303012 CGGTTCTCTGGACAGATACTTCTATCAC
   
 PBH\_0303013 TGCAGTTCCTGGAAGACTCTTCT
   
 PBH\_0303014 CGGTATCTGCACAGAGGCATCT
   
 PBH\_0303015 TGCTGGAAGACGCGTCCA
   
 PBH\_0303016 CGGTTACTGGAGAGACACTTCCATA
   
 PBH\_0303017 CGGCCTGATGAGGAGTACTGG
   
 PBH\_0303018 CCTGTGCGCGAGTCTCGGA
   
 PBH\_0303019 GGCCTGATGCCGAGTACTGG
   
 PBH\_0303020 CAGGAGGAGTCTCTGCGCTT
   
 PBH\_0303021 GAGCAGAAGCGGGGCCGG
   
 PBH\_0303022 TCCTGGAGCGGAGGCCGG
   
 PBH\_0303023 GCGGGCCCTGGTGGACA
   
 PBH\_0303024 GGGGGAGTTCCGGGCGG
   
 PBH\_0303025 GGGGGAGTACCGGGCGG
   
 PBH\_0303026 GGCCTGACGCTGAGTACTGG
   
 PBH\_0303027 CAATGGGACGGAGCGGGTGC
   
 PBH\_0303027a AATGGGACGGAGCGGGTG
   
 PBH\_0303027b GGGACGGAGCGGGT
   
 PBH\_0303028 GGGGGAGTTCGGGGCG
   
 PBH\_0303029 TGGGGGAGTACCGGGCG
   
 PBH\_0303030 ACCAAGAGGAGTACGTGCGCTT
   
 PBH\_0303031 GCCTGCTGCGGAGCACTG
   
 PBH\_0303032 CCAGGAGGAGTTCGTGCGC
   
 PBH\_0303033 CCTGGAAGACGAGCGGGC
   
 PBH\_0303034 GCCTGCTGCGGAGCACTG
   
 PBH\_0303035 GGCCTGATGCCGAGTACTGG
   
 PBH\_0303036 CCAGGAGGAGAACGTGCGC
   
 PBH\_0303037 CCTGGAAGACGAGCGGGC
   
 PBH\_0303038 GACAGGCGCGCCGCG
   
 PBH\_0303039 CTGGAGCAGAGGCGGGC
   
 PBH\_0303040 AACCAAGAGGAGTACGTGCGC
   
 PBH\_0303041 AATGGGACGAGCGGGTG
   
 PBH\_0303055 CATCCTGGAAGACGAGCGGGG

---

# 实施例 4、探针阵列

本实施例提供一种探针阵列。本阵列中有 64 种检测探针，2 种阳性探针，1 种阴性探针，1 种阴性杂交对照探针，1 种阳性杂交对照探针，如表 2 所示，表中，斜体表示各种对照阳性探针，正体表示各种检测探针。

表 2 探针阵列

杂交对照探针 (阳性)	杂交对照探针 (阳性)	检测探针 1	检测探针 1	检测探针 2	检测探针 2	检测探针 3	检测探针 3	检测探针 4	检测探针 4	杂交对照探针 (阳性)	杂交对照探针 (阳性)
杂交对照探针 (阴性)	杂交对照探针 (阴性)	检测探针 5	检测探针 5	检测探针 6	检测探针 6	检测探针 7	检测探针 7	检测探针 8	检测探针 8	杂交对照探针 (阴性)	杂交对照探针 (阴性)
阳性探针 (1)	阳性探针 (1)	检测探针 9	检测探针 9	检测探针 10	检测探针 10	检测探针 11	检测探针 11	检测探针 12	检测探针 12	阳性探针 (1)	阳性探针 (1)
阳性探针 (2)	阳性探针 (2)	检测探针 13	检测探针 13	检测探针 14	检测探针 14	检测探针 15	检测探针 15	检测探针 16	检测探针 16	阳性探针 (2)	阳性探针 (2)
阴性探针	阴性探针	检测探针 17	检测探针 17	检测探针 18	检测探针 18	检测探针 19	检测探针 19	检测探针 20	检测探针 20	阴性探针	阴性探针
检测探针 53	检测探针 53	检测探针 21	检测探针 21	检测探针 22	检测探针 22	检测探针 23	检测探针 23	检测探针 24	检测探针 24	检测探针 57	检测探针 57
检测探针 54	检测探针 54	检测探针 25	检测探针 25	检测探针 26	检测探针 26	检测探针 27	检测探针 27	检测探针 28	检测探针 28	检测探针 58	检测探针 58
检测探针 55	检测探针 55	检测探针 29	检测探针 29	检测探针 30	检测探针 30	检测探针 31	检测探针 31	检测探针 32	检测探针 32	检测探针 59	检测探针 59
检测探针 56	检测探针 56	检测探针 33	检测探针 33	检测探针 34	检测探针 34	检测探针 35	检测探针 35	检测探针 36	检测探针 36	检测探针 60	检测探针 60
阴性探针	阴性探针	检测探针 37	检测探针 37	检测探针 38	检测探针 38	检测探针 39	检测探针 39	检测探针 40	检测探针 40	阴性探针	阴性探针
阳性探针 (2)	阳性探针 (2)	检测探针 41	检测探针 41	检测探针 42	检测探针 42	检测探针 43	检测探针 43	检测探针 44	检测探针 44	阳性探针 (2)	阳性探针 (2)
阳性探针 (1)	阳性探针 (1)	检测探针 45	检测探针 45	检测探针 46	检测探针 46	检测探针 47	检测探针 47	检测探针 48	检测探针 48	阳性探针 (1)	阳性探针 (1)
杂交对照探针 (阳性)	杂交对照探针 (阳性)	检测探针 49	检测探针 49	检测探针 50	检测探针 50	检测探针 51	检测探针 51	检测探针 52	检测探针 52	杂交对照探针 (阳性)	杂交对照探针 (阳性)
杂交对照探针 (阴性)	杂交对照探针 (阴性)	检测探针 61	检测探针 61	检测探针 62	检测探针 62	检测探针 63	检测探针 63	检测探针 64	检测探针 64	杂交对照探针 (阴性)	杂交对照探针 (阴性)

表 3 HLA 分型微阵列中探针的种类和排布

Hex	A001COM	A001CO	A001COM	control	blank	blank	A062-1			IC1	A03Contro	A03Contro	A03Contro
P	NP	P		blank	blank	blank	175	175	175	11	11	11	11
A030				A009						IC2	A03Contro	A03Contro	A03Contro
159a	159a	159a		171a	171a	171a	139	139	139	13	13	13	13
159b	159b	159b		129	129	129	A062-2	139b	139b	139b			
				172	172	172	139a	139a	139a				
A076				A114			140	140	140	A070			
156	156	156		115	115	115	A065			157	157	157	
181	181	181		144	144	144	174a	174a	174a	162	162	162	
182	182	182		114	114	114	174b	174b	174b				
A074				A080			A148			A156			
178	178	178		141a	141a	141a	190	190	190	151b	151b	151b	
180a	180a	180a		141	141	141	192	192	192	152a	152a	152a	
				A111			121a	121a	121a	152b-KC29	152b-KC29	152b-KC29	152b-KC29
				143a	143a	143a	A161			698	698	698	98
				112	112	112	125	125	125				
				A142			153	153	153	A166			
				148b	148b	148b	A163			128	128	128	
IC2				146b	146b	146b	195b	195b	195b	154	154	154	
A03Contro	A03Contro	A03Contro		control	blank	blank				Hex	Hex	Hex	Hex
13	13	13		blank	blank	blank	195a	195a	195a	A001COM	A001COM	A001COM	A001COM
IC1										P	P	P	P
A03Contro	A03Contro	A03Contro											
11	11	11											



## 序列表

&lt;160&gt;214

&lt;210&gt;1

&lt;211&gt;18

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;1

cctgcgctcttggaccgc 18

&lt;210&gt;2

&lt;211&gt;19

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;2

cctgcgctcttggaccgcg 19

&lt;210&gt;3

&lt;211&gt;20

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;





U-15-17

63

<400>3

cctcctgcgctcttggaccg 20

<210>4

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>4

cctgcgctcttggacc 16

<210>5

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>5

cgtgtcccgcccgccg 16

<210>6

<211>16

<212>DNA

<213>人工序列

<220>



<223>

<400>6

atggagccgcgggcgc 16

<210>7

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>7

cctgcgtcttggaccgcgg 20

<210>8

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>8

gcggtccaagagcgcagg 18

<210>9

<211>18

<212>DNA

<213>人工序列



<220>

<223>

<400>9

cctgcgcttttggaccgc 18

<210>10

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>10

cctgcgctgttggaccgc 18

<210>11

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>11

gcaggagaggcctgagtattgg 22

<210>12

<211>24

<212>DNA

<213>人工序列



<220>

&lt;223&gt;

<400>12

caccatcagataatgtatggctgc 24

&lt;210&gt;13

&lt;211&gt;25

&lt;212&gt;DNA

### 〈213〉人工序列

 $\langle 220 \rangle$ 

&lt;223&gt;

<400>13

caccatccagataatgtatggctgc 25

&lt;210&gt;14

 $\langle 211 \rangle_{20}$ 

## <212>DNA

### 〈213〉人工序列

 $\langle 220 \rangle$ 

&lt;223&gt;

<400>14

ttctacacctccgtgtcccg 20

&lt;210&gt;15

&lt;211&gt;19

## <212>DNA

  
<213>人工序列

<220>

<223>

<400>15

cgettcacgcagtgaggct 19

<210>16

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>16

cgagccagaagatggagcc 19

<210>17

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>17

ccgcgggcaccgtggata 18

<210>18



<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>18

gcaggagggtccggagtatt 20

<210>19

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>19

gacgtggggccggacggg 18

<210>20

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>20

gacgggcgcctcctccgc 18



<210>21

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>21

cggggtaccaccagtagcct 20

<210>22

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>22

ggtaccggcaggacgccta 19

<210>23

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>23

cgccctgaacgaggacctg 19



<210>24

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>24

cggacatggcagctcagatc 20

<210>25

<211>

<212>DNA

<213>人工序列

<220>

<223>

<400>25

ccaccaagcacaaagtggga 19

<210>26

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>26





aagtgggagacggcccatg 19

<210>27

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>27

aggcggcccgtgtggcgg 18

<210>28

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>28

aggcgggtccatgcggcgg 18

<210>29

<211>18

<212>DNA

<213>人工序列

<220>

<223>



<400>29

cggcccatgaggcggagc 18

<210>30

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>30

tacctggatggcacgtgcg 19

<210>31

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>31

ctggaggcgagtgcggtgg 19

<210>32

<211>18

<212>DNA

<213>人工序列

<220>

<223>



<400>32

tgcggtggacgggctccgc 18

<210>33

<211>24

<212>DNA

<213>人工序列

<220>

<223>

<400>33

gtattttctacacctccgtgtcccg 24

<210>34

<211>

<212>DNA

<213>人工序列

<220>

<223>

<400>34

cgagcggtttgacagcgac 19

<210>35

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>35

cgtgcgggttcgacagcgac 19

<210>36

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>36

cgtggggccggacggg 16

<210>37

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>37

aggcgggtccatgcggcg 17

<210>38

<211>18

<212>DNA

<213>人工序列



<220>

<223>

<400>38

cccggccgcggggagccc 18

<210>39

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>39

ccgcgggcgccgtggata 18

<210>40

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>40

tgggacgaggagacaggga 19

<210>41

<211>19

<212>DNA

<213>人工序列



07.06.20

51

<220>

<223>

<400>41

tgggaccaggagacacgga 19

<210>42

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>42

tggggaccctgcgcggcta 19

<210>43

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>43

gacgtggggtcggacggg 18

<210>44

<211>18

<212>DNA



<213>人工序列

<220>

<223>

<400>44

gacgggcgcttcctccgc 18

<210>45

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>45

gcgggtaccagcaggacgc 19

<210>46

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>46

cgccctgaaagaggacctg 19

<210>47

<211>20



&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;47

agctcagatcaccaagcgca 20

&lt;210&gt;48

&lt;211&gt;21

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;48

tcagatcaccaagcgcaagag 21

&lt;210&gt;49

&lt;211&gt;20

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;49

agctcagatcaccgagcgca 20

&lt;210&gt;50

&lt;211&gt;20





<212>DNA

<213>人工序列

<220>

<223>

<400>50

ggctcagatcaccacgca 20

<210>51

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>51

tcagatcaccacgcaagtg 21

<210>52

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>52

agacggcccatgagcg 17

<210>53



<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>53

agacggcccatgaggcgg 18

<210>54

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>54

gcgggagcagcggagagtct 19

<210>55

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>55

agacggcccatgaggcgg 18



<210>56

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>56

gcggagcagttgagagcct 19

<210>57

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>57

ggcggagcagttgagagcc 19

<210>58

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>58

gcggagcagtgagagcct 19



<210>59

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>59

tacctggagggcacgtgcg 19

<210>60

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>60

tgcgtagtggtccgc 18

<210>61

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>61



tcaccgagtggaacctgggg 19

<210>62

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>62

ccgagtggaacctggggacc 19

<210>63

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>63

tgaccgagagaacctgcgg 19

<210>64

<211>20

<212>DNA

<213>人工序列

<220>

<223>



<400>64

ccgagagaacctgcggatcg 20

<210>65

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>65

gaaggcccactcacagactg 20

<210>66

<211>23

<212>DNA

<213>人工序列

<220>

<223>

<400>66

tattttttcacatccgtgtcccg 23

<210>67

<211>21

<212>DNA

<213>人工序列

<220>

<223>



03-03-07

6

<400>67

tctacacttcggtttcccggc 21

<210>68

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>68

ctacacctccatgtcccggc 20

<210>69

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>69

ccggaacacacggaagttaa 21

<210>70

<211>19

<212>DNA

<213>人工序列

<220>



00:00:00

61

<400>70

attgggacggggagacacg 19

<210>71

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>71

gacacggaatatgaaggcca 21

<210>72

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>72

gacacggaatgtgaaggccc 20

<210>73

<211>22

<212>DNA

<213>人工序列





00-05-30

1)

<220>

<223>

<400>73

tcacagactcaccgagtgacc 22

<210>74

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>74

tcacagattgaccgagtgacc 22

<210>75

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>75

tcacagactgaccgagtgacc 22

<210>76

<211>18

<212>DNA

<213>人工序列



03-05-30

68

<220>

<223>

<400>76

cgagcgaacctggggacc 18

<210>77

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>77

ccgagagagcctgcggatc 19

<210>78

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>78

accgagagaacctggggacc 20

<210>79

<211>19

<212>DNA



07-05-00

69

<213>人工序列

<220>

<223>

<400>79

gtggacctggcgaccctgc 19

<210>80

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>80

caccgtccagaggatgtatggc 22

<210>81

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>81

accagcaggacgcttacgacg 21

<210>82

<211>20



00-05-30

)<sup>6</sup>

<212>DNA

<213>人工序列

<220>

<223>

<400>82

tcgccttgaacgaggacctg 20

<210>83

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>83

cctgcgctcttggaccgc 18

<210>84

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>84

tcagaccaccaagcacaagtgg 22

<210>85



03 05-30

71

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>85

gaggcggcccatgtggc 17

<210>86

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>86

ggcccatgcggcggagc 17

<210>87

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>87

gcggcccgctcggcgga 17



00:00:00

72

<210>88

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>88

gcacgtgcgtggagtggc 18

<210>89

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>89

gccggtgcgtggacgggc 18

<210>90

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>90

ggcgagtgctggagtggc 19



0005-10

7}

<210>91

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>91

gcacgtgcgtggacgggc 18

<210>92

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>92

gccggtgcgtggagtggc 18

<210>93

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>93



05-05-2004

74

ggcggagtgcgtggacgggc 19

<210>94

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>94

agacacggaaagtgaaggccc 21

<210>95

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>95

tggccctgaccgagacctgggc 22

<210>96

<211>20

<212>DNA

<213>人工序列

<220>

<223>





03-05-20

15

<400>96

ctacaaccagagcgaggccg 20

<210>97

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>97

gccctgacccagacctggg 19

<210>98

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>98

cccgaaacctcctcctgc 18

<210>99

<211>18

<212>DNA

<213>人工序列

<220>

<223>



000000

76

<400>99

cccgaaccgtcctcctgc 18

<210>100

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>100

tgctctcggcgccctg 17

<210>101

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>101

tgctctcgggagccctgg 18

<210>102

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>102

ggggggcagtgccct 16

<210>103

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>103

tgaggatatttcgacaccgcca 21

<210>104

<211>23

<212>DNA

<213>人工序列

<220>

<223>

<400>104

tgaggatatttctacaccgcatg 23

<210>105

<211>20

<212>DNA

<213>人工序列



&lt;220&gt;

&lt;223&gt;

&lt;400&gt;105

tttccacacctccgtgtccc 20

&lt;210&gt;106

&lt;211&gt;19

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;106

tctacaccgccatgtccc 19

&lt;210&gt;107

&lt;211&gt;20

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;107

tctacacctccgtgtcccg 20

&lt;210&gt;108

&lt;211&gt;22

&lt;212&gt;DNA

&lt;213&gt;人工序列



<220>

<223>

<400>108

ccgcttcacatcagtgaggctac 22

<210>109

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>109

cgcttcacacccgtgggct 19

<210>110

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>110

cgcttcacgcagtgaggct 19

<210>111

<211>20

<212>DNA



0255 00

80

<213>人工序列

<220>

<223>

<400>111

tacgtggacggcaccagtt 20

<210>112

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>112

cgtggacgacaccagttcg 20

<210>113

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>113

ggacgacacgctgttcgtga 20

<210>114

<211>20



81

<212>DNA

<213>人工序列

<220>

<223>

<400>114

tggacgacacgcagttcgtg 20

<210>115

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>115

gcgacgccacgagtcg 17

<210>116

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>116

gcgacgccgcgagtc 16

<210>117



82

<211>

<212>DNA

<213>人工序列

<220>

<223>

<400>117

gagtccgagagaggagccgc 20

<210>118

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>118

ccgaggaaggagccgcg 17

<210>119

<211>15

<212>DNA

<213>人工序列

<220>

<223>

<400>119

aggatggcgccccgg 15





83

&lt;210&gt;120

&lt;211&gt;16

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;120

ggacggagcccccgggc 16

&lt;210&gt;121

&lt;211&gt;18

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;121

cgggcgccgtggatagag 18

&lt;210&gt;122

&lt;211&gt;18

&lt;212&gt;DNA

&lt;213&gt;人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt;122

cgggcgccatggatagag 18



000000

84

<210>123

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>123

ggggccggaatattgggac 19

<210>124

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>124

ggggccggagtattgggac 19

<210>125

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>125



03 05-00

85

gggaccgggagacacagatct 21

<210>126

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>126

tgggaccggaacacacagatc 21

<210>127

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>127

acacagaagtacaagcgccagg 22

<210>128

<211>21

<212>DNA

<213>人工序列

<220>

<223>



07-07-30

86

<400>128

acacggaacatgaaggcctcc 21

<210>129

<211>24

<212>DNA

<213>人工序列

<220>

<223>

<400>129

cacacagatcttcaagaccaacac 24

<210>130

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>130

atctgcaaggccaaggcaca 20

<210>131

<211>21

<212>DNA

<213>人工序列

<220>

<223>



03.05.20

81

<400>131

tacaaggcccaggcacagact 21

<210>132

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>132

acacagactgaccgagag 18

<210>133

<211>23

<212>DNA

<213>人工序列

<220>

<223>

<400>133

cacacagacttaccgagagagcc 23

<210>134

<211>16

<212>DNA

<213>人工序列

<220>



03.06.00

88

<223>

<400>134

gcaccgcgctccgcta 16

<210>135

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>135

cggaccctgctccgctact 19

<210>136

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>136

acctgcggatcgcgctc 17

<210>137

<211>16

<212>DNA

<213>人工序列



03-07-20

sf

<220>

<223>

<400>137

cggaacctgcgcggct 16

<210>138

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>138

cgggtctcacatcatccagagg 22

<210>139

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>139

gggtctcacaccctccagagg 21

<210>140

<211>23

<212>DNA

<213>人工序列



07 06 00

9c

<220>

<223>

<400>140

tcacacttggcagacgatgtatg 23

<210>141

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>141

acaccctccagaggatgtacgg 22

<210>142

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>142

cgacctggggcccgac 16

<210>143

<211>16

<212>DNA





930100

8

<213>人工序列

<220>

<223>

<400>143

cgacgtggggccggac 16

<210>144

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>144

gggtaccaccaggacgcct 19

<210>145

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>145

cgggtatgaccaggacgcc 19

<210>146

<211>18



<212>DNA

<213>人工序列

<220>

<223>

<400>146

gggcatgaccagtcgcc 18

<210>147

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>147

gcgggtataaccagttcgcc 20

<210>148

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>148

gaggacctgcgctcctgga 19

<210>149



<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>149

gaggacctgagctcctgga 19

<210>150

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>150

ggaccgccgcggacac 16

<210>151

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>151

ggaccgccgcggacac

16



<210>152

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>152

cggacacggcggtcag 17

<210>153

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>153

cggacaccgcggtcag 17

<210>154

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>154

ggcccgtaggcggag 16



<210>155

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>155

ggcccggtgtggcggag 16

<210>156

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>156

gcggagcaggacagagccta 20

<210>157

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>157



03-06-00

9/

gcggagcagtgagagccta 20

<210>158

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>158

gcggagcagctgagagccta 20

<210>159

<211>23

<212>DNA

<213>人工序列

<220>

<223>

<400>159

agcagctgagaacctacctggag 23

<210>160

<211>23

<212>DNA

<213>人工序列

<220>

<223>



03-05-200

9

<400>160

agcagctgagagcctacctggag 23

<210>161

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>161

ggagggcgagtgcggtgg 17

<210>162

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>162

ggagggcacgtgcgtgg 17

<210>163

<211>17

<212>DNA

<213>人工序列

<220>

<223>



03 05 20

68

<400>163

ggagggcctgtgcgtgg 17

<210>164

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>164

cgtggagtcgctccgcag 18

<210>165

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>165

cgtggagtggtccgcag 18

<210>166

<211>21

<212>DNA

<213>人工序列

<220>





03.05.20

88

<223>

<400>166

ctccgcagacacctggagaac 21

<210>167

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>167

gctccgcagatacctggagaa 21

<210>168

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>168

aggacaagctggagcgcg 18

<210>169

<211>17

<212>DNA

<213>人工序列



03-05-20

10-4

<220>

<223>

<400>169

ggacacgctggagcgcg 17

<210>170

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>170

ggagacgctgcagcgcg 17

<210>171

<211>25

<212>DNA

<213>人工序列

<220>

<223>

<400>171

cttgtggcagcttaagtttgaatgt 25

<210>172

<211>25

<212>DNA

<213>人工序列



101

101

<220>

<223>

<400>172

tggagtactctacgtctgagtgta 25

<210>173

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>173

ggagcaggttaaacatgagtg 22

<210>174

<211>23

<212>DNA

<213>人工序列

<220>

<223>

<400>174

cctgtggcagggttaagtataagt 23

<210>175

<211>23

<212>DNA



000000

10}

<213>人工序列

<220>

<223>

<400>175

ttggagtactctacgggtgagtg 23

<210>176

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>176

cctgtggcagcctaagaggg 20

<210>177

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>177

cctggagcaggcgcg 16

<210>178

<211>18



<212>DNA

### <213>人工序列

<220>

&lt;223&gt;

&lt;400&gt;178

cctggaagacgagcgggc 18

&lt;210&gt;179

&lt;211&gt;19

&lt;212&gt;DNA

### 〈213〉人工序列

<220>

&lt;223&gt;

&lt;400&gt;179

ccaggaggagaacgtgcgc 19

 $\langle 210 \rangle_{180}$ 

<211>18

&lt;212&gt;DNA

### 〈213〉人工序列

&lt;220&gt;

&lt;223&gt;

 $\langle 400 \rangle 180$ 

cctggaagacaggcgggc 18

&lt;210&gt;181



03.01.00

14/2

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>181

cggttgctggaaagatgcac 21

<210>182

<211>27

<212>DNA

<213>人工序列

<220>

<223>

<400>182

cggttcctggacagatacttctatcac 27

<210>183

<211>24

<212>DNA

<213>人工序列

<220>

<223>

<400>183

tgcagttcctggaaagactcttct 24



<210>184

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>184

cggtatctgcacagaggcatct 22

<210>185

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>185

tgctggaagacgcgtcca 19

<210>186

<211>25

<212>DNA

<213>人工序列

<220>

<223>

<400>186

cggttactggagagacacttccata 25



030107

1/1

<210>187

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>187

cggcctgatgaggagtactgg 21

<210>188

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>188

cctgtcgccgagtcctgga 19

<210>189

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>189





03 03 03

(10)

ggcctgatgccgagtactgg 20

<210>190

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>190

caggaggagctcctgcgctt 20

<210>191

<211>

<212>DNA

<213>人工序列

<220>

<223>

<400>191

gagcagaagcggggccgg 18

<210>192

<211>17

<212>DNA

<213>人工序列

<220>

<223>



03 06 30

608

<400>192

tcctggagcggaggcgg 17

<210>193

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>193

gcgggcccctggtggaca 17

<210>194

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>194

gggggagttccgggcgg 17

<210>195

<211>17

<212>DNA

<213>人工序列

<220>

<223>



000000

10

<400>195

ggggggagtaccggggcgg 17

<210>196

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>196

ggcctgacgctgagtactgg 20

<210>197

<211>20

<212>DNA

<213>人工序列

<220>

<223>

<400>197

caatgggacggagcgggtgc 20

<210>198

<211>18

<212>DNA

<213>人工序列

<220>



03.01.00

11

<223>

<400>198

aatgggacggagcgggtg 18

<210>199

<211>14

<212>DNA

<213>人工序列

<220>

<223>

<400>199

gggacggagcgggt 14

<210>200

<211>16

<212>DNA

<213>人工序列

<220>

<223>

<400>200

gggggagttccgggcg 16

<210>201

<211>17

<212>DNA

<213>人工序列



00 00 00

(((

<220>

<223>

<400>201

tgggggagtaccgggcg 17

<210>202

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>202

accaagaggagtacgtgcgctt 22

<210>203

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>203

gcctgctgcggagcactg 18

<210>204

<211>19

<212>DNA

<213>人工序列



170800

(1)

<220>

<223>

<400>204

ccaggaggagttcgtgcgc 19

<210>205

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>205

cctggaagacgagcgggc 18

<210>206

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>206

gcctgctgcggagcactg 18

<210>207

<211>20

<212>DNA



00 00 00

11}

<213>人工序列

<220>

<223>

<400>207

ggcctgatgccgagtactgg 20

<210>208

<211>19

<212>DNA

<213>人工序列

<220>

<223>

<400>208

ccaggaggagaaacgtgcgc 19

<210>209

<211>18

<212>DNA

<213>人工序列

<220>

<223>

<400>209

cctggaagacgagcgggc 18

<210>210

<211>15



<212>DNA

<213>人工序列

<220>

<223>

<400>210

gacaggcgcgcgcgcg 15

<210>211

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>211

ctggagcagaggcgggc 17

<210>212

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>212

aaccaagaggagtacgtgcgc 21

<210>213





00 00 00

105

<211>17

<212>DNA

<213>人工序列

<220>

<223>

<400>213

aatgggacgcagcggBt 17

<210>214

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>214

catcctggaagacgagcgggg 21



03-05-30

161

# 说明书附图



图 1



图 2

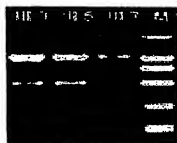


图 3



图 4



05-05-20

117

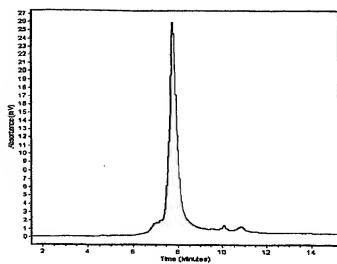


图 5A

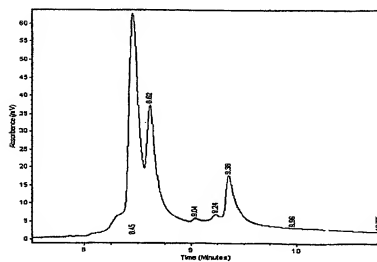


图 5B



03-05-00

68

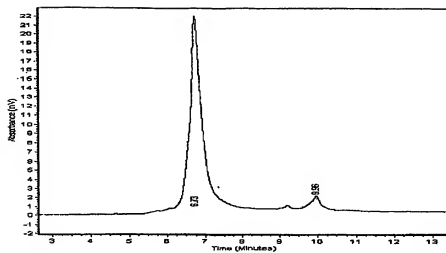


图 6A

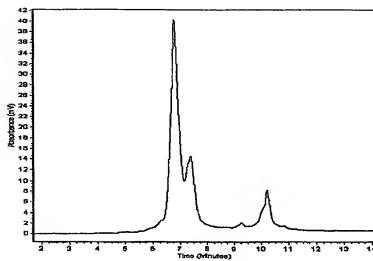


图 6B



03-06-20

12

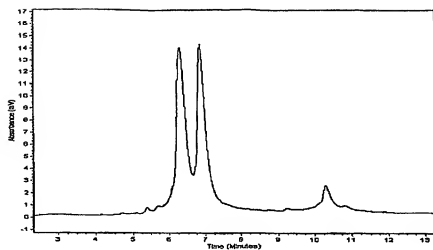


图 6C

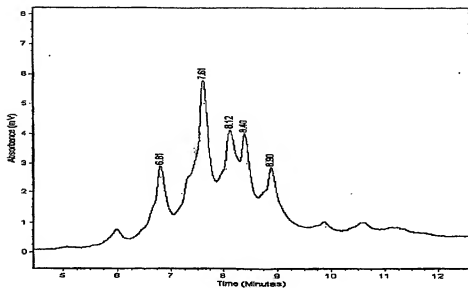


图 6D

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☒ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**